GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

GM protein protein search, using sw model

Run on: November 21, 2003, [illegible] ; Search time 45.5 seconds
(without alignments)
8819.4 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEEENIPA[illegible]EIHAASSRSTGMQEE...SII 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 66718 seqs, 18259940 residues

Total number of hits satisfying chosen parameters: 66718

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US08A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09A_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US09B_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score listed. The smaller the Pred. No., the more likely the result is to be true and not by chance. For single-sequence searches, Pred. No. is derived by analysis of the statistical significance of the high-scoring tail of the score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8853	100.0	1715	12	US-10-021-571-2	Sequence 2, Appli
2	8853	100.0	1715	12	US-10-117-029-6	Sequence 6, Appli
3	8839.5	99.8	1762	12	US-10-117-029-7	Sequence 7, Appli
4	8266	93.4	1715	12	US-10-021-571-4	Sequence 4, Appli
5	8246	93.1	1715	12	US-10-117-029-2	Sequence 2, Appli
6	8212	92.8	1771	12	US-10-117-029-4	Sequence 4, Appli
7	8212	92.8	1771	12	US-10-117-029-11	Sequence 11, Appl
8	5930.5	67.0	1184	12	US-10-117-029-9	Sequence 9, Appli
9	5741.5	64.7	1184	12	US-10-117-029-9	Sequence 9, Appli
10	2961.5	33.5	705	15	US-[illegible]-638-[illegible]	Sequence 6[illegible], Ap
11	2782.5	31.4	551	9	US-09-841-788A-17	Sequence 17, Appl
12	2782.5	31.4	551	12	US-10-117-042-17	Sequence 17, Appl
13	2776.5	31.3	1498	12	US-10-021-571-6	Sequence 6, Appli
14	2260	25.5	513	15	US-[illegible]-819-6	Sequence 6, Appli
15	1848.5	20.9	1399	12	US-10-021-571-6	Sequence 6, Appli
16	549	6.2	109	11	US-09-986-480-371	Sequence 371, App
17	538.5	6.1	1724	10	US-09-964-899-43	Sequence 43, Appl
18	483.5	5.5	1762	12	US-10-205-194-117	Sequence 117, App
19	438	4.9	740	9	US-09-839-788A-12	Sequence 12, Appl
20	438	4.9	740	12	US-10-175-042-12	Sequence 12, Appl
21	401.5	4.5	426	9	US-09-908-711-70	Sequence 70, Appl
22	395.5	4.5	1327	9	US-09-841-835-2	Sequence 2, Appli
23	395.5	4.5	1327	11	US-09-972-115A-8	Sequence 8, Appli
24	395.5	4.5	1327	12	US-10-199-937-4	Sequence 4, Appli
25	394.5	4.5	765	12	US-10-128-174-3	Sequence 3, Appli
26	394.5	4.5	765	12	US-10-128-174-34	Sequence 34, Appl
27	394.5	4.5	765	12	US-10-128-174-35	Sequence 35, Appl
28	394.5	4.5	765	12	US-10-128-174-36	Sequence 36, Appl
29	394.5	4.5	765	12	US-10-128-174-37	Sequence 37, Appl
30	394.5	4.5	765	12	US-10-128-174-38	Sequence 38, Appl
31	394.5	4.5	765	12	US-10-128-174-39	Sequence 39, Appl
32	394.5	4.5	765	12	US-10-128-174-40	Sequence 40, Appl
33	394.5	4.5	765	12	US-10-128-174-41	Sequence 41, Appl
34	394.5	4.5	765	12	US-10-128-174-42	Sequence 42, Appl
35	394.5	4.5	765	12	US-10-128-174-43	Sequence 43, Appl
36	394.5	4.5	765	12	US-10-128-174-44	Sequence 44, Appl
37	394	4.5	1619	12	US-10-369-978-4	Sequence 4, Appli
38	392	4.4	1554	12	US-10-301-822-93	Sequence 93, Appl
39	392	4.4	1704	12	US-10-369-978-2	Sequence 2, Appli
40	389.5	4.4	285	9	US-09-839-788A-18	Sequence 18, Appl
41	389.5	4.4	285	12	US-10-175-042-18	Sequence 18, Appl
42	387.5	4.4	949	9	US-09-841-835-10	Sequence 10, Appl
43	381.5	4.3	1181	12	US-10-199-937-139	Sequence 139, App
44	380.5	4.3	1166	11	US-09-972-115A-6	Sequence 6, Appli
45	380.5	4.3	1166	12	US-10-199-937-135	Sequence 135, App

ALIGNMENTS

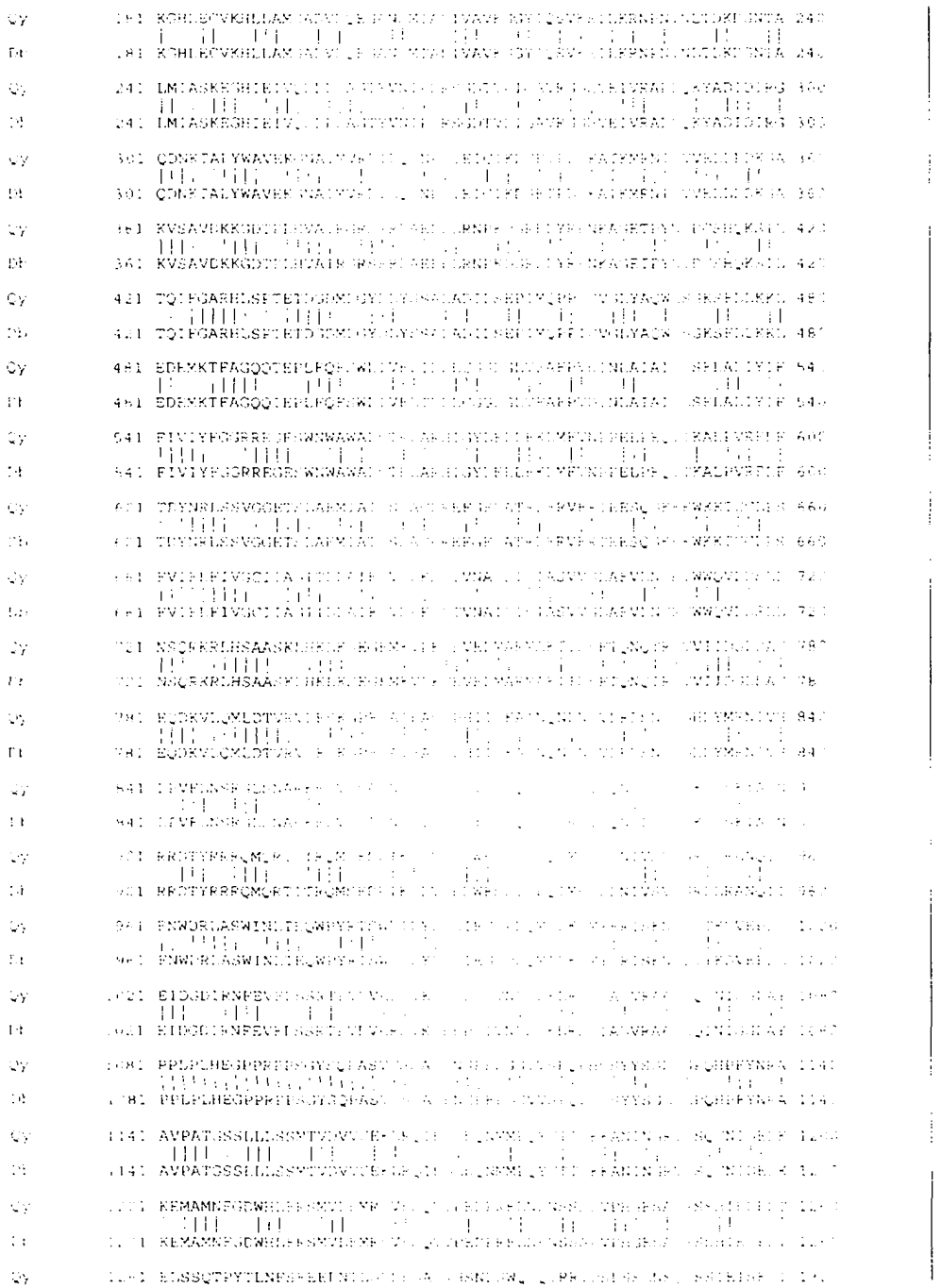
RESULT 1
US-10-021-571-2
; Sequence 2, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Huiyong
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL AN
; FILE REFERENCE: CHA 1.1A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-021-571-2

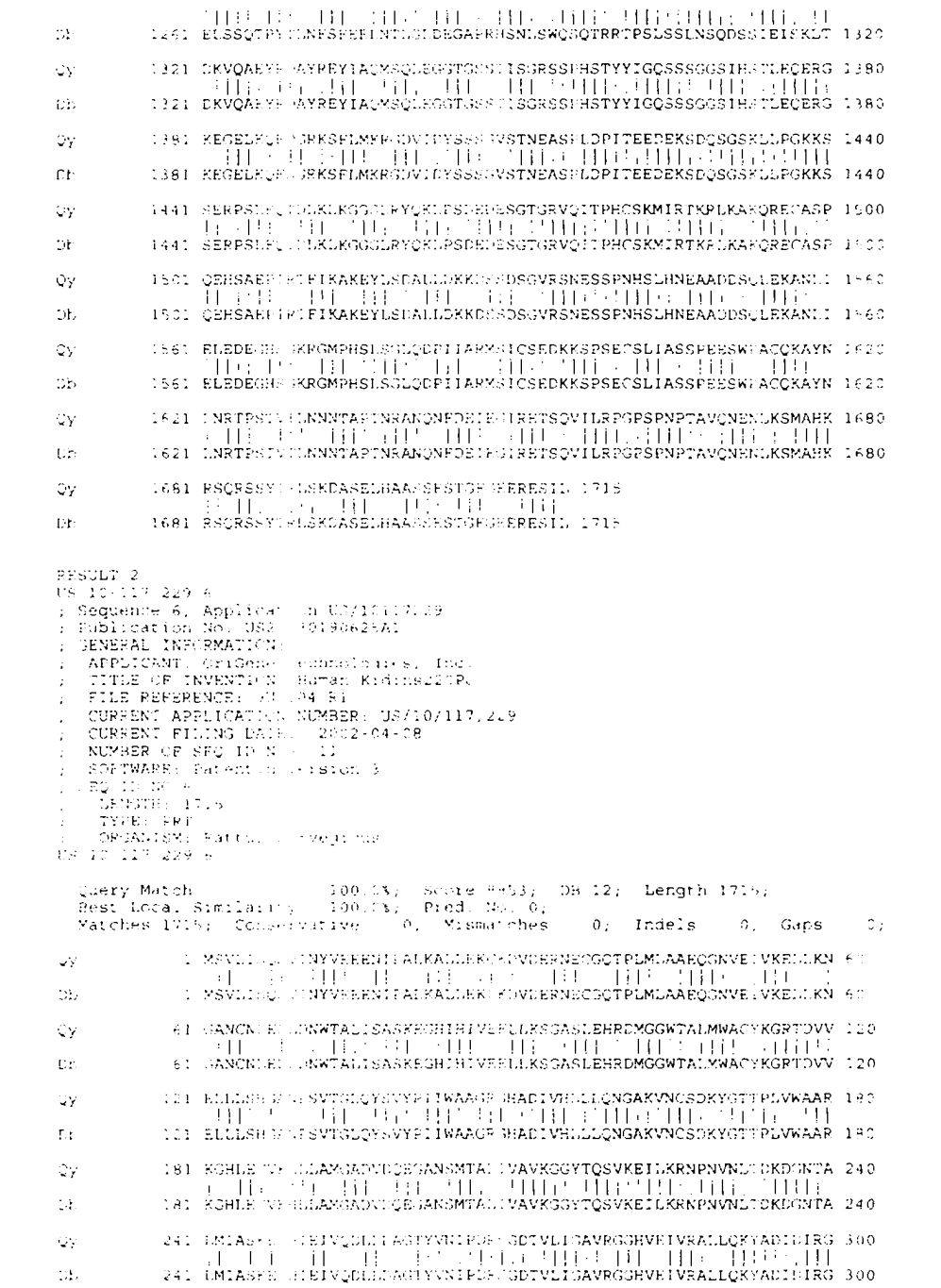
Query Match 100.0%; Score 8853; DB 12; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60
Db 1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60

Qy 61 GANCNLEDLCNWTALISASKEGHIHIVEFLLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
Db 61 GANCNLEDLCNWTALISASKEGHIHIVEFLLKSGASLEHRDMGGWTALMWACYKGRTDVV 120

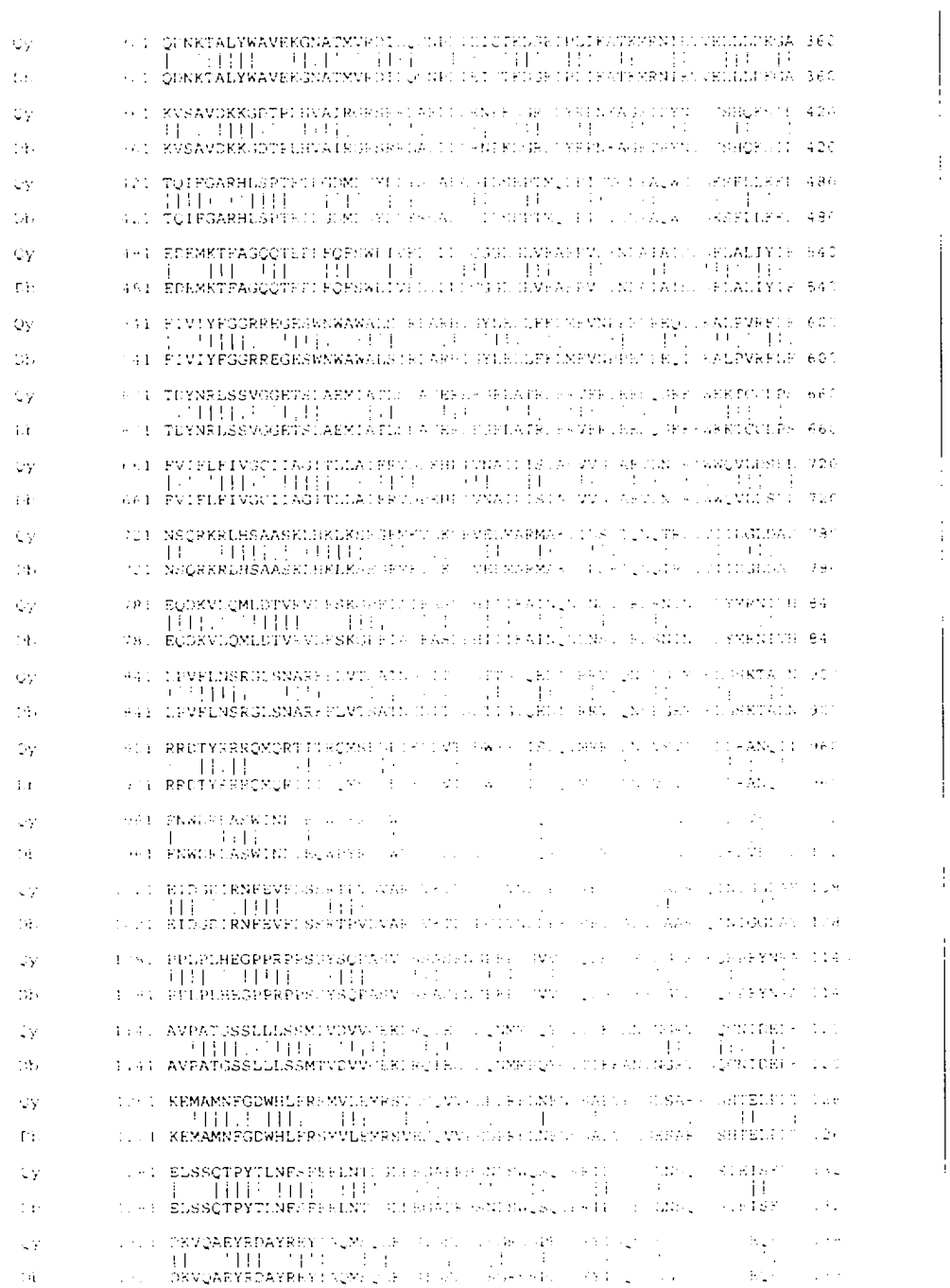
Qy 121 ELLLSHGANVSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANVSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

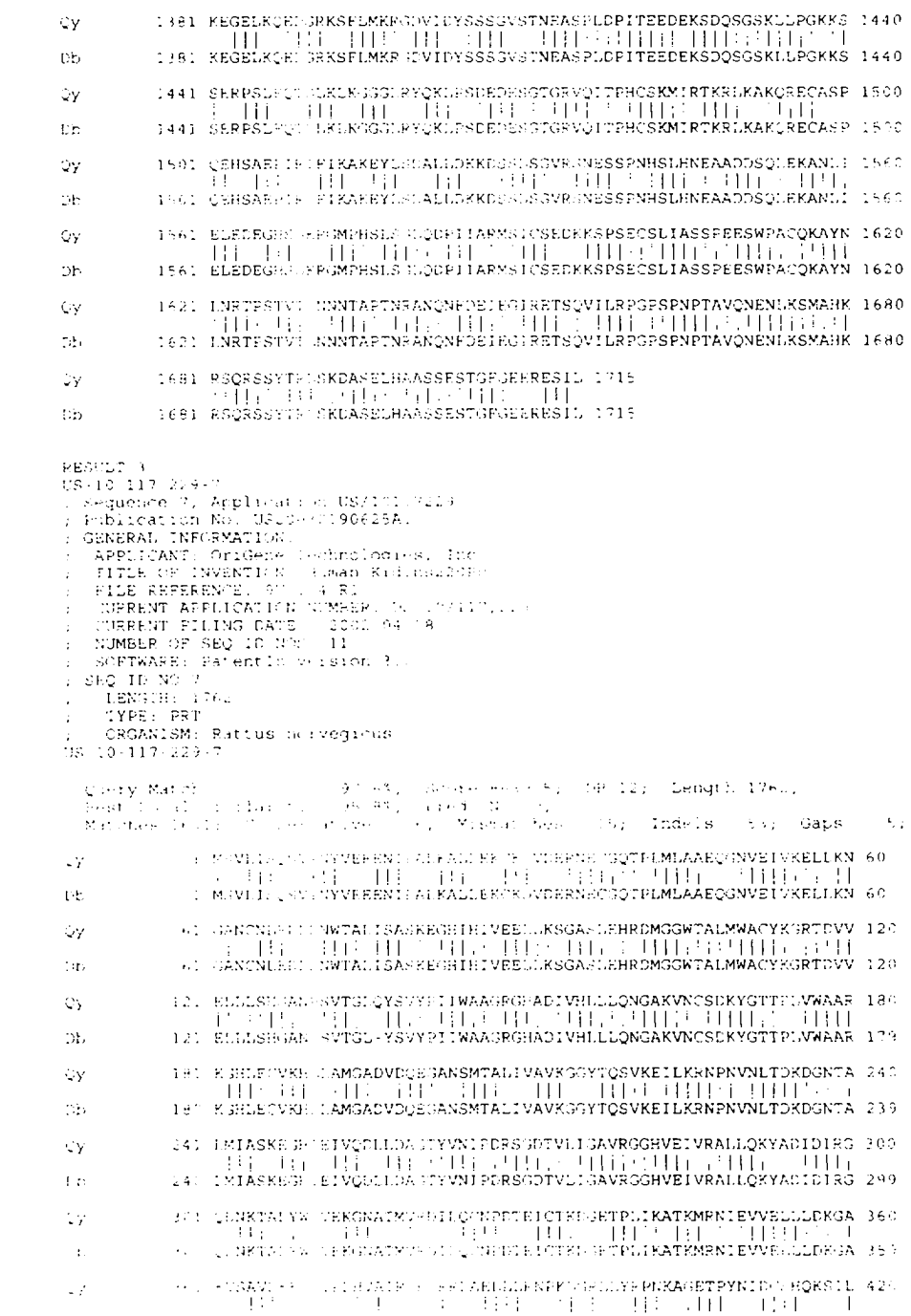




RESULT 2
US 10/117,229-6
; Sequence 6, Applicat n US/10/117,229
; Publication No. US2 0196625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidney cDNA
; FILE REFERENCE: 47 04 P1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID N : 11
; SOFTWARE: PatentIn version 3
; SEQ ID NO 6
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US 10/117,229-6

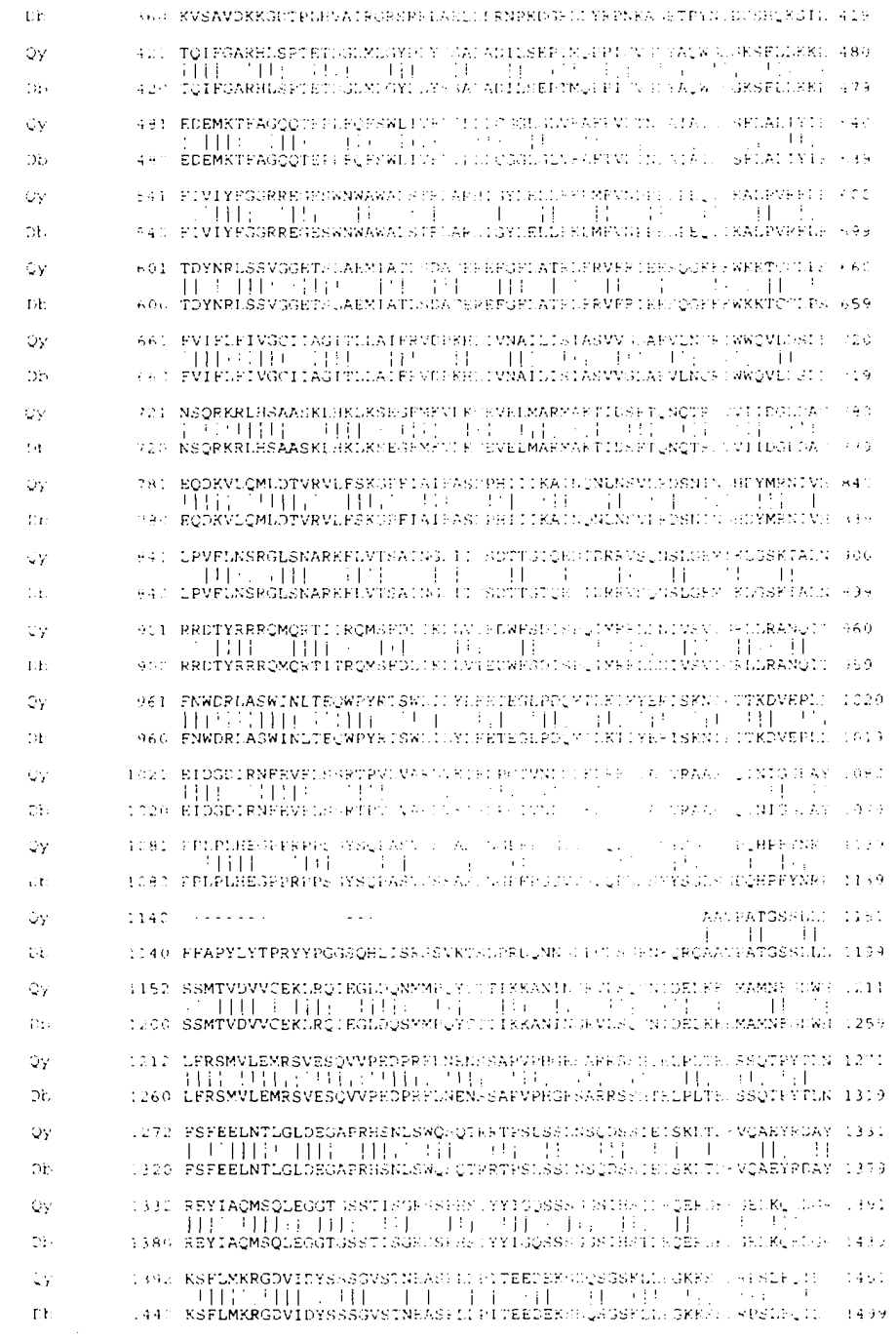
Query Match 100.0%; Score 8953; DB 12; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0, Mismatches 0; Indels 0, Gaps 0;

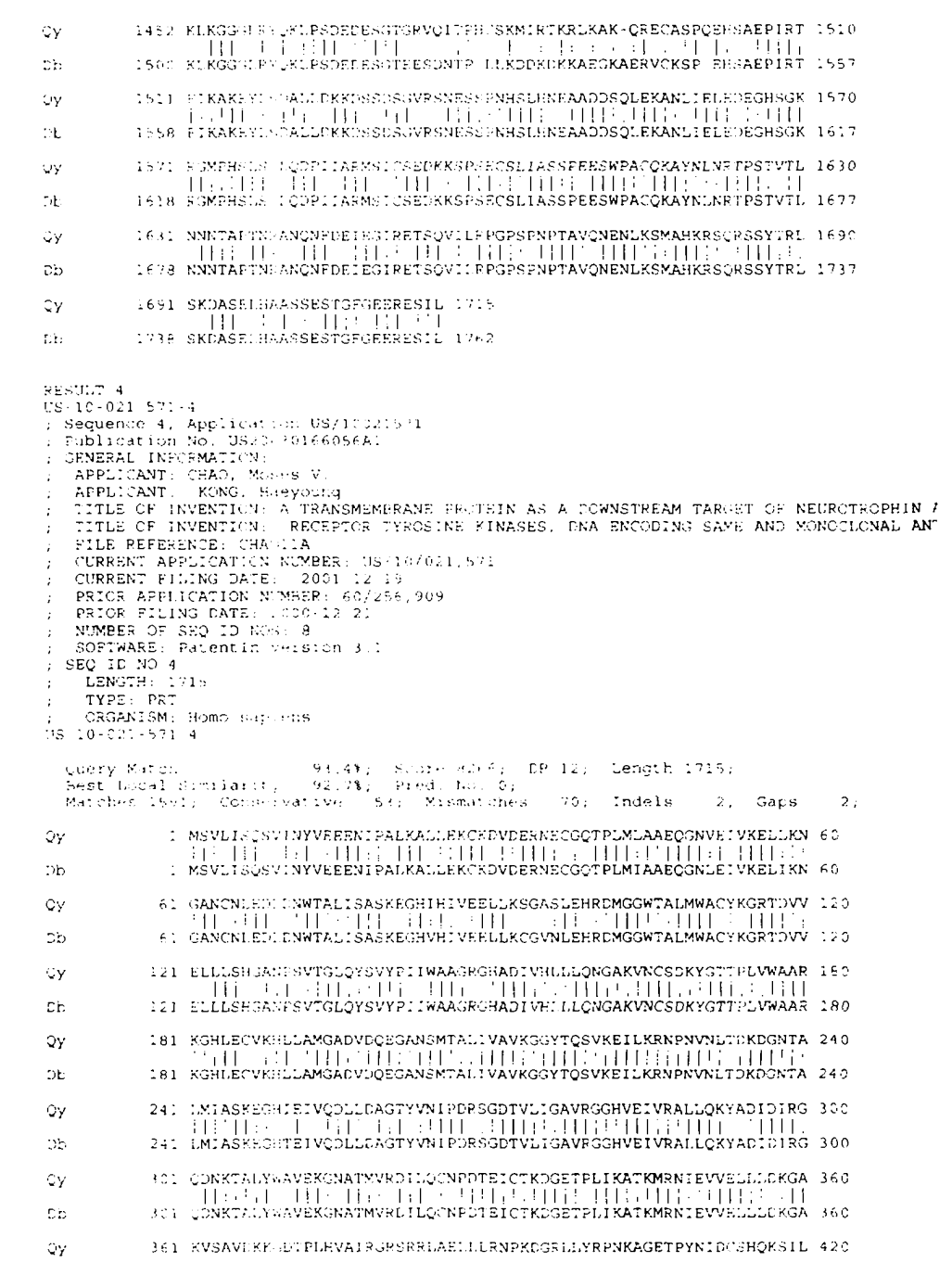




RESULT 3
US-10-117-229-7
; Sequence 7, Application US/10117229
; Publication No. US2003190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Human Kinases
; FILE REFERENCE: 97.4 B
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 1764
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-7

Query Match 97.4%; Score 5; DB 12; Length 1764;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1740; Conservative 5; Mismatches 5; Indels 5; Gaps 5;





RESULT 4
US-10-021-571-4
; Sequence 4, Application: US/10/021,571
; Publication No. US20030166066A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Hueyjung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN /
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL AN
; FILE REFERENCE: CHAO-11A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-571-4

Query Match 91.4%; Score 4064; DP 12; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 54; Mismatches 70; Indels 2; Gaps 2;

Qy 1 MSVLIRQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60
Db 1 MSVLIRQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Qy 61 GANCNLEDCDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
Db 61 GANCNLEDCDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

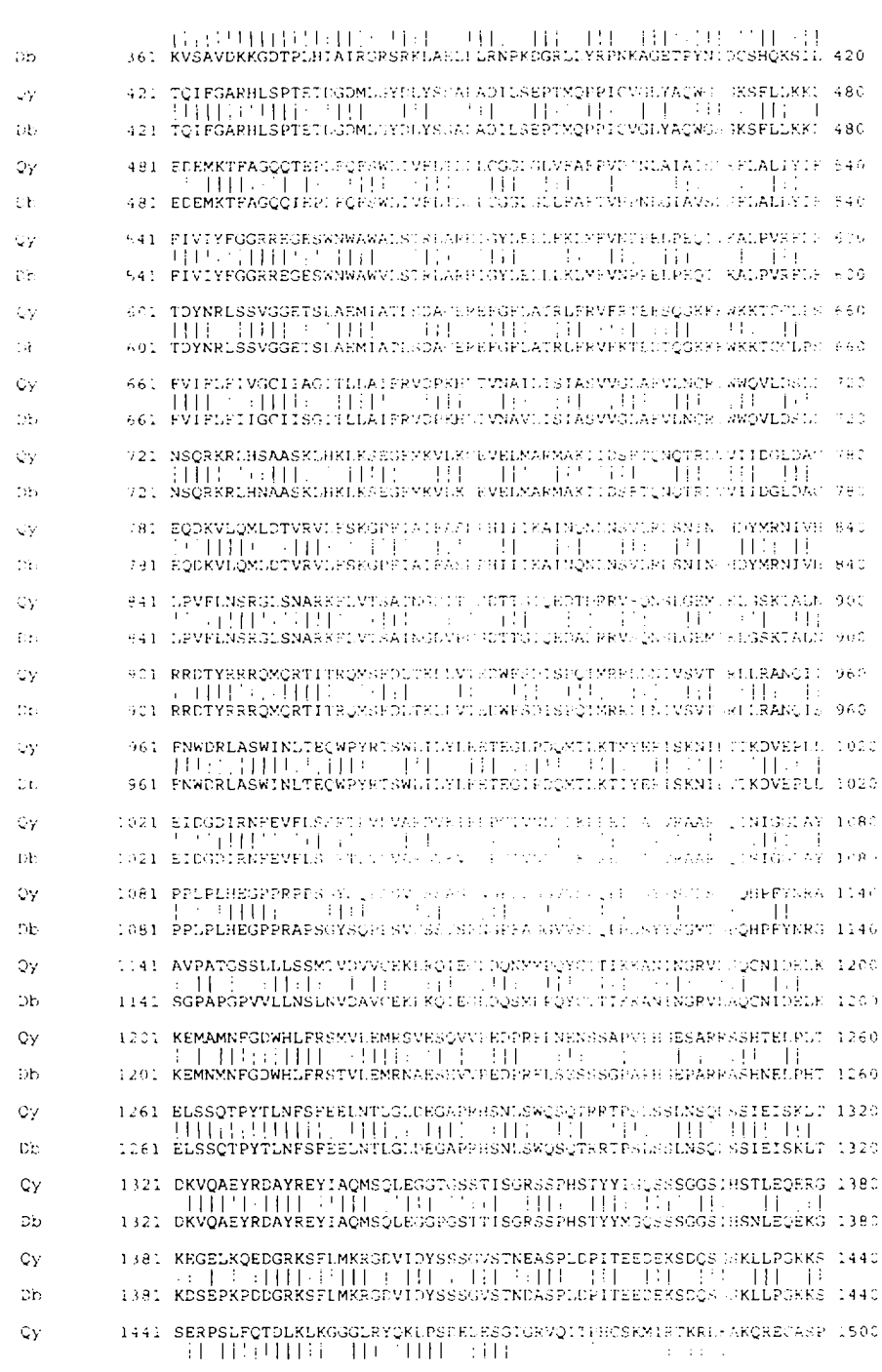
Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

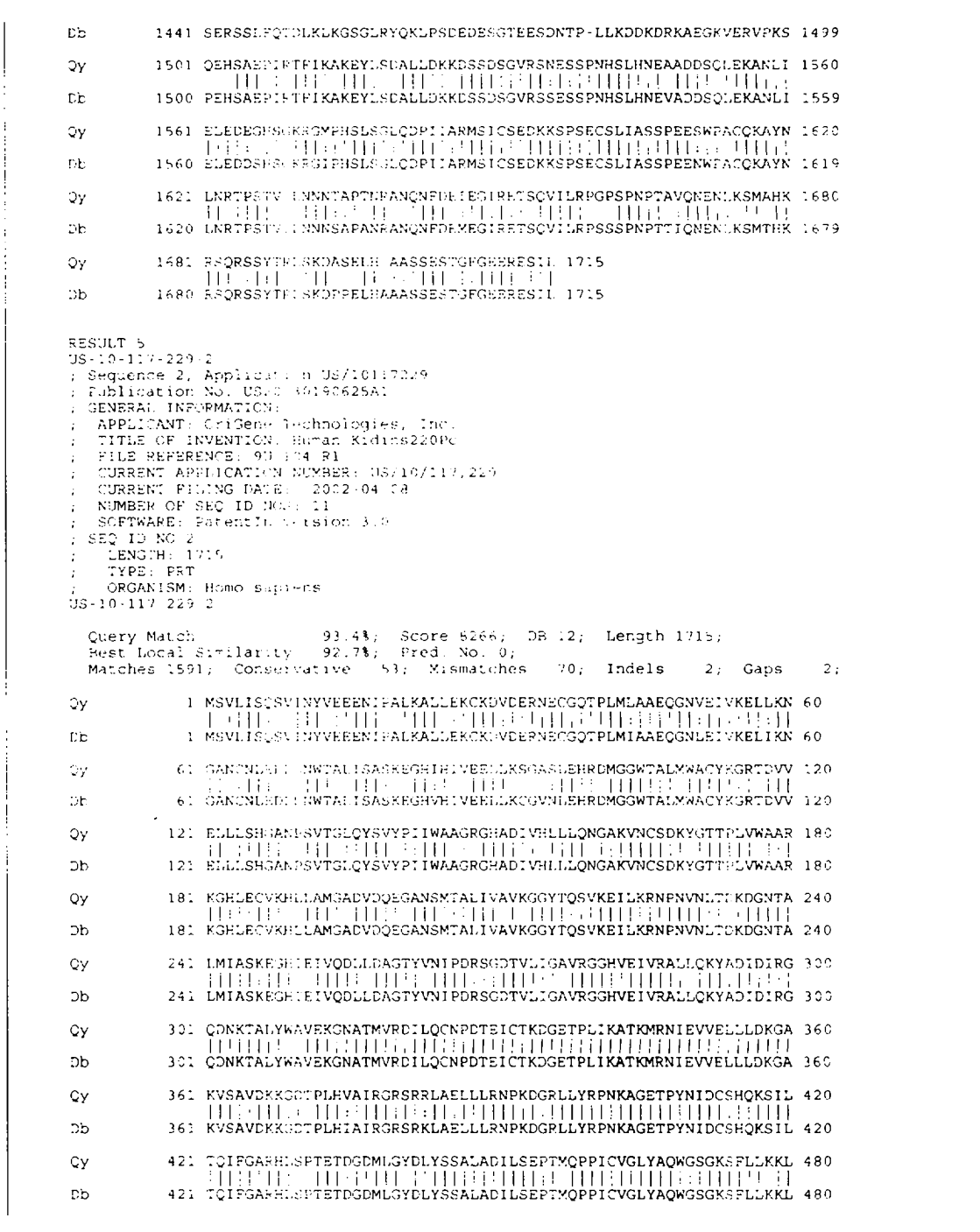
Qy 181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
Db 181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy 241 LMIASKEGHIEIVQDLLDAGTYVNIPDPSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
Db 241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy 301 QDNKTALYWAVEKGNATMVRDILQQNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
Db 301 QDNKTALYWAVEKGNATMVRDILQQNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

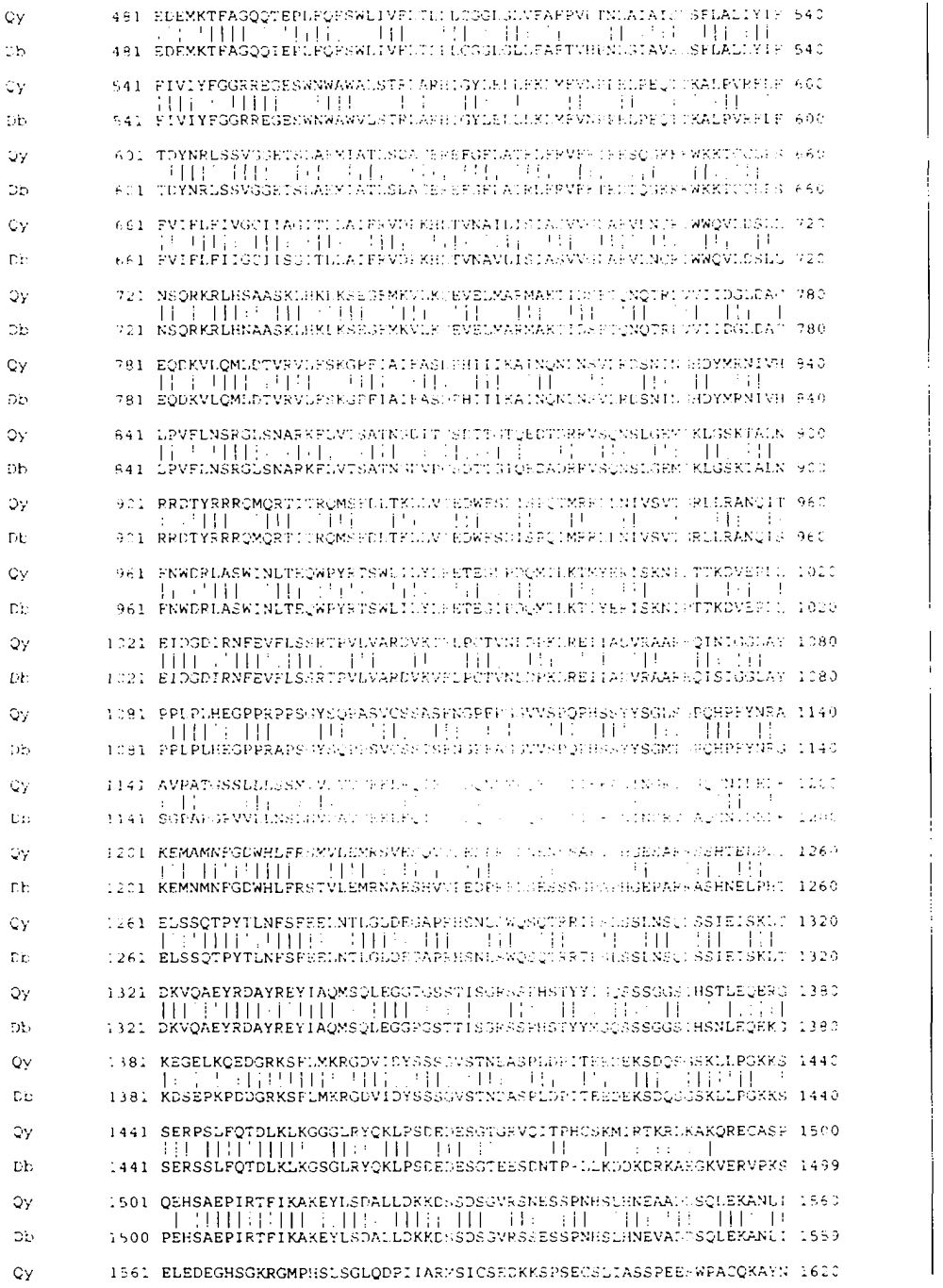
Qy 361 KVSAVIKKEDTPLHVAIRGRSRRIAELLLRNPKDGRLLYRPNKAGETPYNIDQSHQKSIL 420





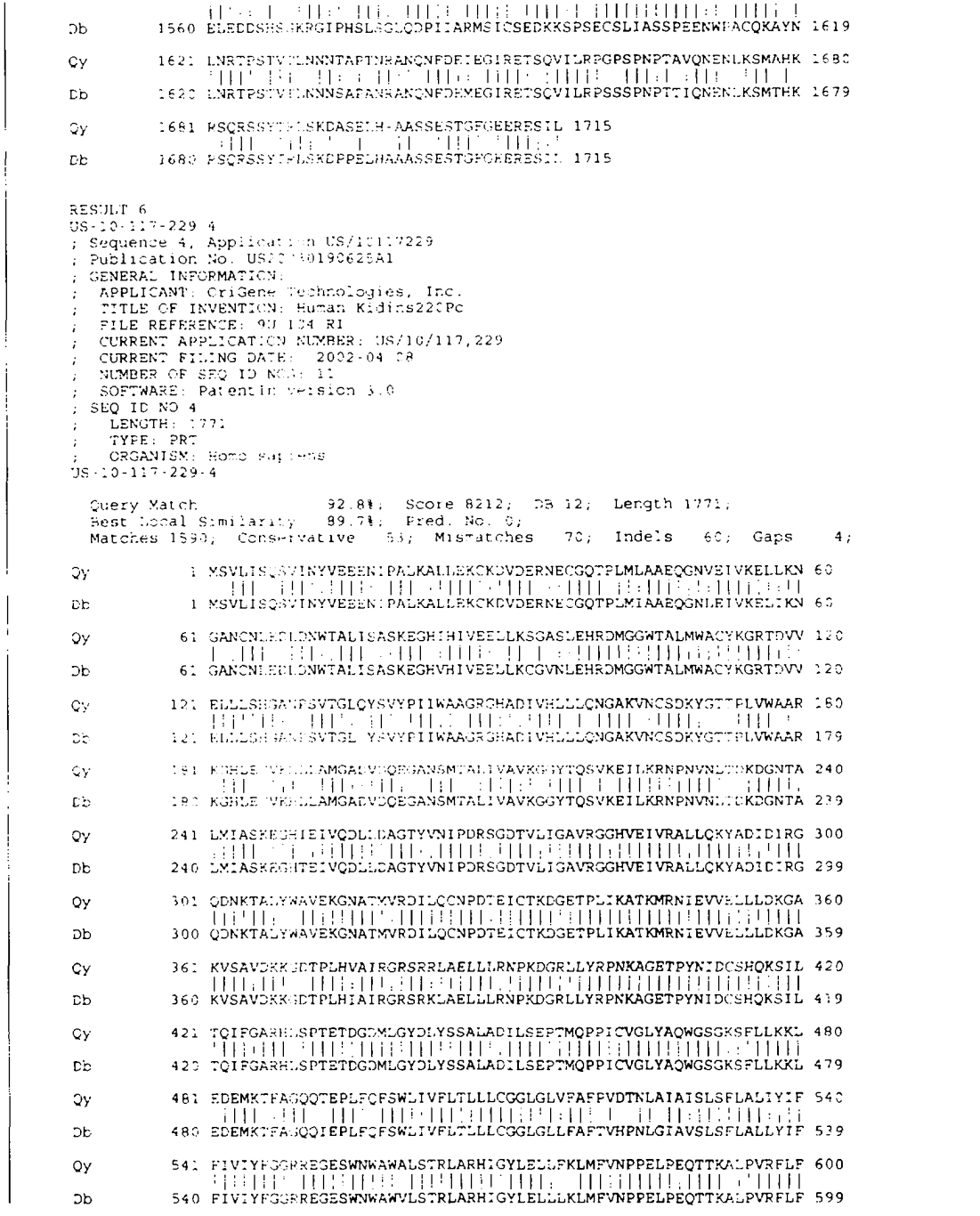
RESULT 5
US-10-117-229-2
; Sequence 2, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: CuraGen Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 90 P34 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-2

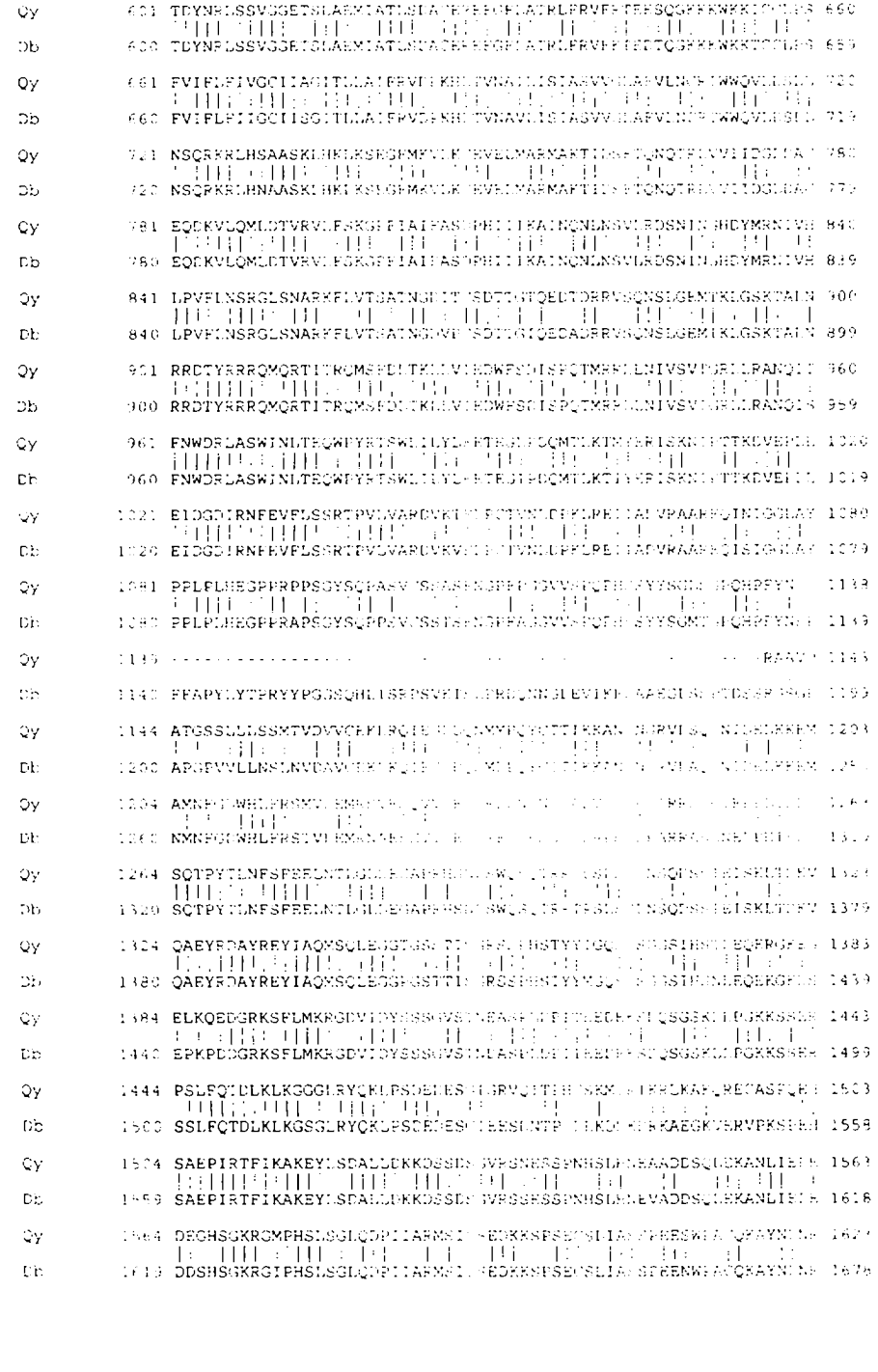
Query Match 93.4%; Score 6266; DB 12; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

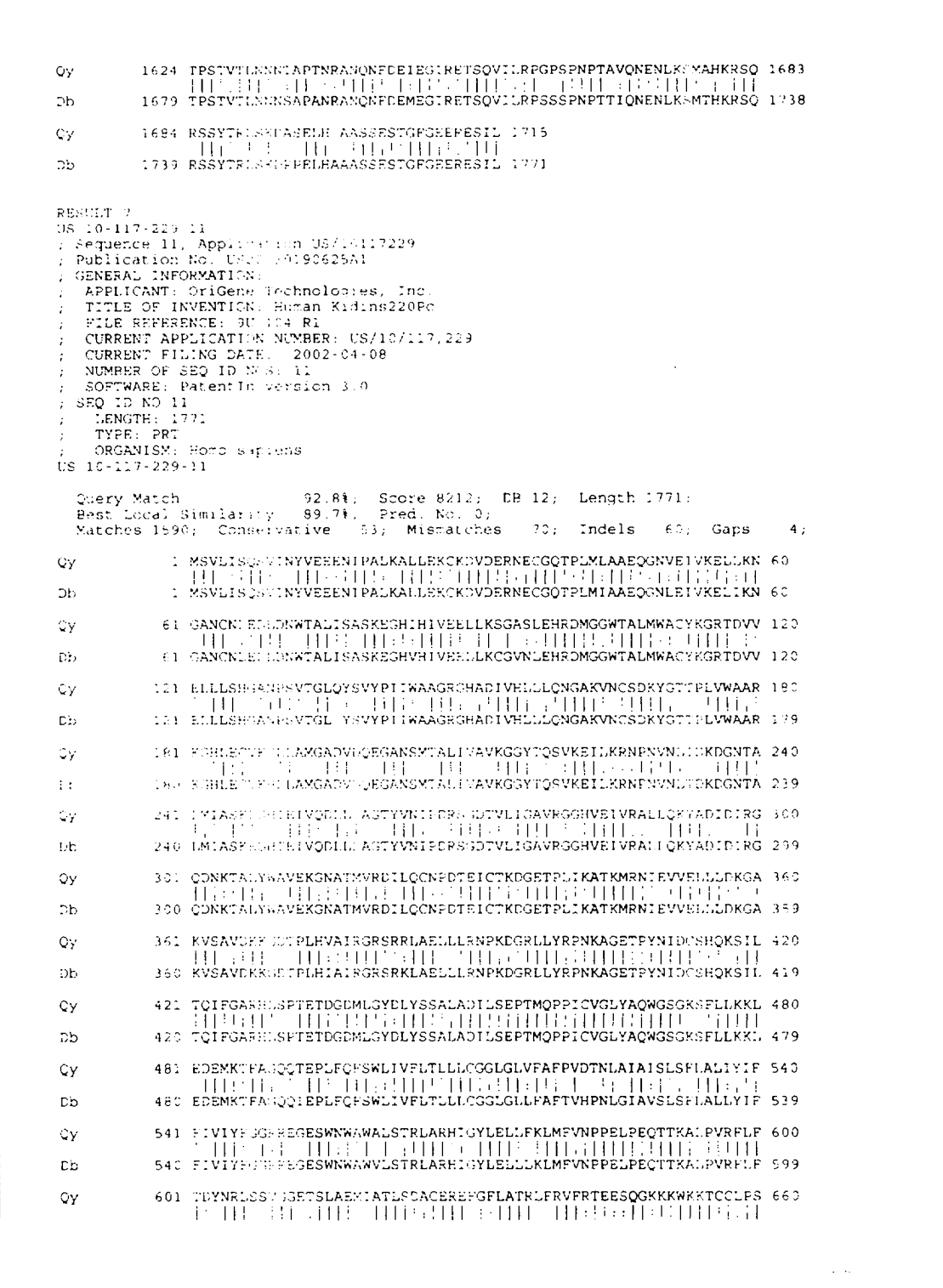


RESULT 6
US-10-117-229 4
; Sequence 4, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: CriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 RI
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04 08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-4

Query Match 92.8%; Score 8212; DB 12; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;

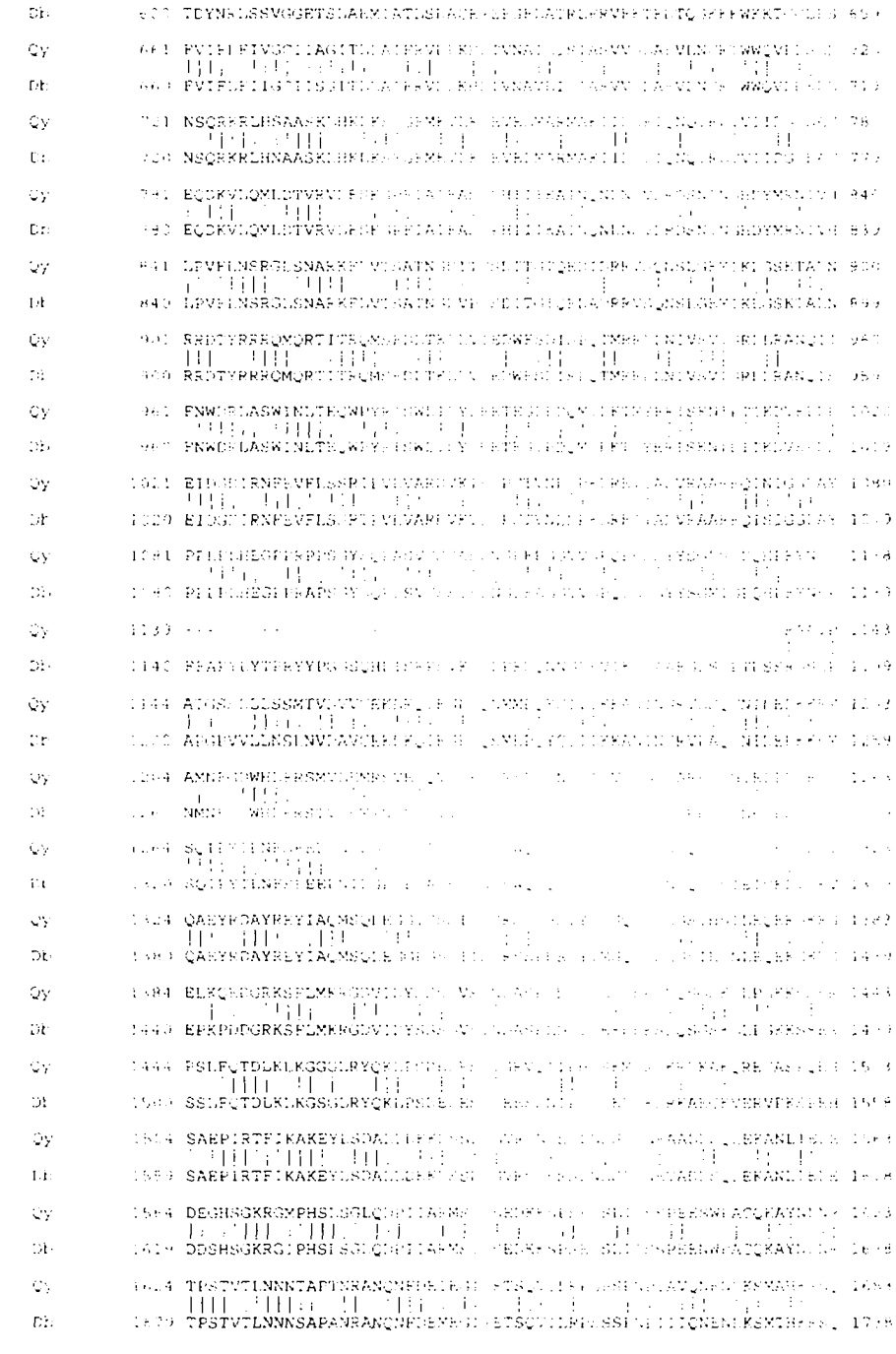






RESULT 2
US 10-117-229-11
; Sequence 11, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 3U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US 10-117-229-11

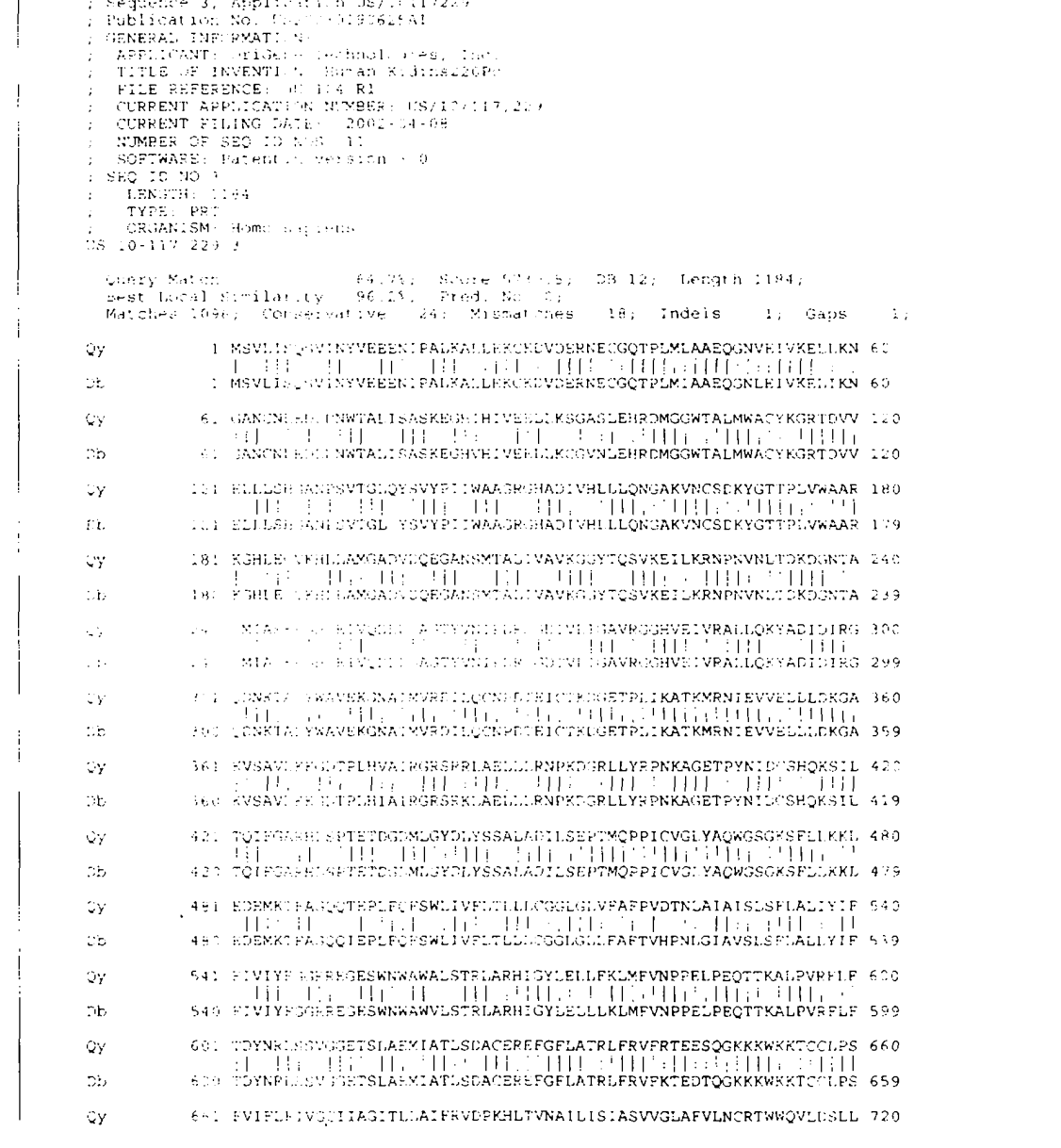
Query Match 92.8%; Score 8212; DB 12; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 83; Mismatches 70; Indels 60; Gaps 4;

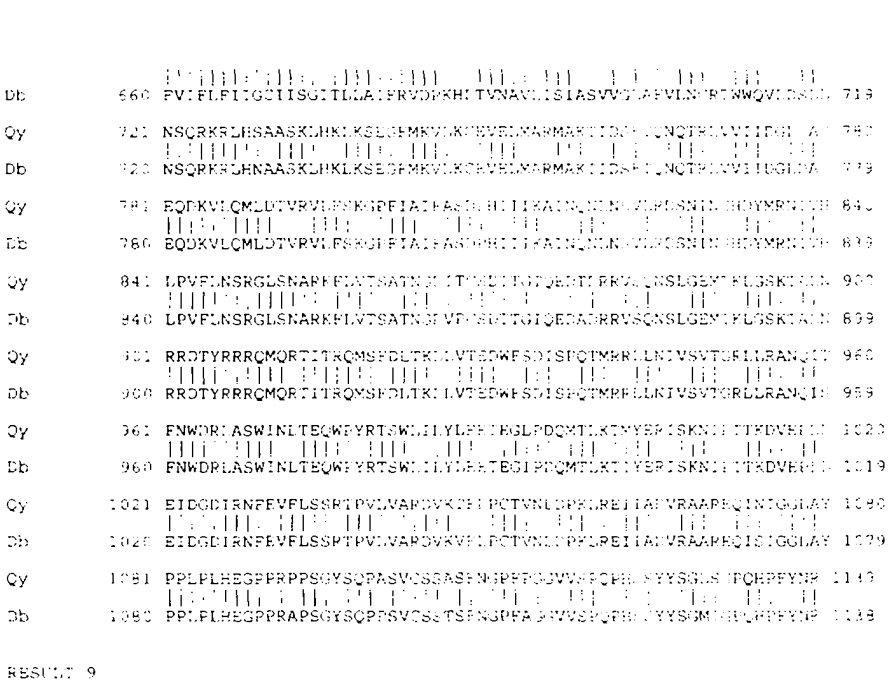


Qy 1684 RSSYTPLSFDASELHAASSESTGFGEERESIL 1715
Db 1739 RSSYTPLSFPPELHAAASSESTGFGEERESIL 1771

RESULT 8
US-10-117-229-3
; Sequence 3, Application US/10117229
; Publication No. [redacted]
; GENERAL INFORMATION:
; APPLICANT: [redacted]
; TITLE OF INVENTION: Human K.[illegible]
; FILE REFERENCE: [redacted]
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent[illegible] Version 3.0
; SEQ ID NO 3
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-3

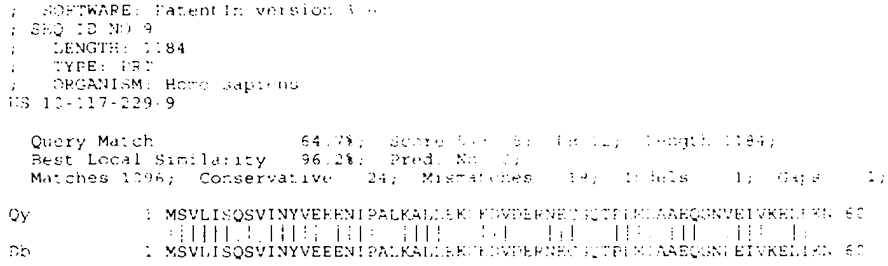
Query Match 66.7%; Score 577.5; DB 12; Length 1184;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

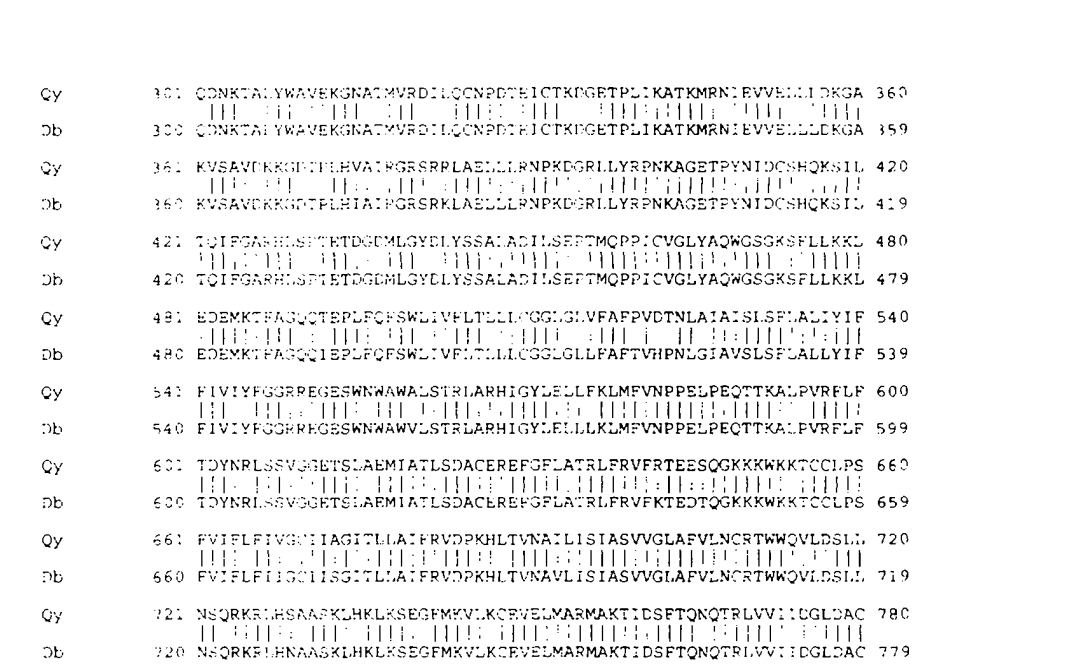




RESULT 9
US-10-117-229-9
; Sequence 9, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220P
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: Homo sapiens
US 10-117-229-9

Query Match 64.7%; Score [illegible]; DB [illegible]; Length 1184;
Best Local Similarity 96.2%; Pred. No. [illegible];
Matches 1096; Conservative 24; Mismatches 14; Indels 1; Gaps 1;





RESULT 10
US-10-106-698-6378
; Sequence 6378, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptid
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280

; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6378
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (244)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (337)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L amino acids
US 10-106-696-6378

Query Match 33.5%; Score 2041.5; DB 15; Length 705;
Best Local Similarity 95.3%; Pred. No. 2e-218;
Matches 567; Conservative 11; Mismatches 14; Indels 1; Gaps 1;

Qy 62 ANCNCLEDLDNWTAICSASKEGIHICYQDCIPGLMSLEHRKCWEHVKEQVCTLKLWDTDVFGTLRLVNSLQVEGTCKR 141
Db 28 SNCNCIEDLDNWTAICSASKEGIHICYQDCIPGLMSLEHRKCWEHVKEQVCTLKLWDTDVFGTLRLVNSLQVEGTCKR 107

Qy 142 LLLSHGANPSVTDLYSVVTCWAAKHCAIVHLLIQNGAPVDCQSEPSRTPLVWAAAKG 201
Db 108 LLLSHGANPSVTDLYSVVTCWAAKHCAIVHLLIQNGAPVDCQSEPSRTPLVWAAAKG 167

Qy 202 GHLECVKHLLAMGADVNQKMNTALYQAVQEKKHEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 261
Db 168 GHLECVKHLLAMGADVDQKMNTALYQAVQEKKHEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 227

Qy 262 MIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 341
Db 228 MIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 307

Qy 342 DNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLHIAIRGRSRKLAELLLRNPKDGRLDVYRP 401
Db 308 DNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLHIAIRGRSRKLAELLLRNPKDGRLDVYRP 367

Qy 402 VSAVIKKGGTPLHLAIRGRSRKLAELLLRNPKDGRLDVYRPNRAGETPYNICSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 481
Db 368 VSAVIKKGGTPLHLAIRGRSRKLAELLLRNPKDGRLDVYRPNRAGETPYNICSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 447

Qy 482 QIPHAPHLSHTEPEIQMALSLTSDAPPKALTMQIAIANGIPPLTTVHPNGNHETLVLLFNHIPVLNVDDQMGETALHCAIRNDSSAAVVLLENGADVNAQ 581
Db 448 QIPHAPHLSHTEPEIQMALSLTSDAPPKALTMQIAIANGIPPLTTVHPNGNHETLVLLFNHIPVLNVDDQMGETALHCAIRNDSSAAVVLLENGADVNAQ 547

Qy 582 DEMKPFAGQQIEPLFQPSWIIVFGLRRFLGSWNWAWALSTPLARHIGYLELLPKLMPFVPPEIPGQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 661
Db 548 DEMKPFAGQQIEPLFQPSWIIVFGLRRFLGSWNWAWALSTPLARHIGYLELLPKLMPFVPPEIPGQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 627

Qy 662 IVIYFGGRREGESWNWAWALSTRLARHIGYLELLPKLMPFVPPEIPGQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 741
Db 628 IVIYFGGRREGESWNWAWVLSTRLARHIGYLELLPKLMPFVPPEIPGQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 707

Qy 742 DYNRLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVPEIPGQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 821
Db 708 DYNRLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVPEIPGQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 787

RESULT 11
US 09-835-788A-17
; Sequence 17, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,595
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US 09-835-788A-17

Query Match 31.4%; Score 2780.5; DB 9; Length 551;
Best Local Similarity 96.7%; Pred. No. 7.5e-205;
Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

Qy 102 MGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNG 161
Db 1 MGGWTALMWACYKGRTDVVELLLSHGANPSVTGL-YSVYPIIWAAGRGHADIVHLLLQNG 59

Qy 162 AKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVK 221
Db 60 AKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVK 119

Qy 222 EILKRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 281
Db 120 EILKRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 179

Qy 282 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 341
Db 180 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 239

Qy 342 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLDVYRP 401
Db 240 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLDVYRP 299

Qy 402 NKAGETPYNICSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 461
Db 300 NKAGETPYNICSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 359

Qy 462 CVGLYAQWGSGKSFLLKKLEDEMKTFAGQQTEPLFQPSWLIVFLTLLLCGGLGLVFAFPV 521
Db 360 CVGLYAQWGSGKSPLLKKLEDEMKTFAGQQIEPLFQPSWLIVFLTLLLCGGLGLLFAFTV 419

Qy 522 FINLAIAKQLA SALIYIPFIVYFQGRFFGESWNWAWALSTPLARHIGYLELLPKLMFV 581
Db 420 FFNLGIAVSLSPLALLYIPFIVYFQGRFFGESWNWAWVLSTRLARHIGYLELLLKLMFV 479

Qy 582 NPPELPEQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 641
Db 480 NPPELPEQTTKALPVRFLPTDYNPLSSVDGETSIAEMIATLSDACEREPGPLATRLPRVP 539

Qy 642 RTEESQDKKKWK 653
Db 540 KTEDTQDKKKKK 551

RESULT 12
US 10-175-042-17
; Sequence 17, Application US/10175042
; Publication No. US20030181710A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides,
; TITLE OF INVENTION: Polypeptides, and
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/10/175,042
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 09/835,788
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666

```
;  PRIOR FILING DATE: 2000-10-17
;  PRIOR APPLICATION NUMBER: 60/[illegible],585
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/[illegible],246
;  PRIOR FILING DATE: 1999-11-24
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;    LENGTH: 551
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-[illegible]-042-17

  Query Match             31.4%;  Score 2382.5;  DB 12;  Length 551;
  Best Local Similarity   96.7%;  Pred. No. 2.1e-[illegible];
  Matches  534;  Conservative    4;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        122 MGGWTALMWACYKGRTDVVBLLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLIQNG 181
              ||||||||.|.||||||.|  |||| |.|||  ||   |.|  |||.|||  |||  |||
Db          1 MGGWTALMWACYKGRTDVVELLLSHGANPSVTGLLYSVYPIIWAAGRGHADIVHLLIQNG  59

Qy        182 AKVNCSDKYGTTPLVWAARKGHLEVVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVE 241
              |||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 AKVNCSDKYGTTPIVWAARKGHLEVVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVE 119

Qy        242 EILKRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 301
              |||||||| ||||| |||| |||| |  |||  ||| |||  |||  |||  |||  | |
Db        120 EILKRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 179

Qy        302 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQKNPDTEITTKDGETPLI 361
              ||| |||| |||| |||| |||| |  ||| |||  |||  |||  |||  |||  |||
Db        180 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQKNPDTEITTKDGETPLI 239

Qy        362 KATKMRNIEVVELLLKKGAKVSAVDKKGDTPLHVAIERGRSRFLAELLLRNPKDGRLLYRN 421
              ||||| |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        240 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHIAIERGRSRFLAELLLRNPKDGRLLYRN 299

Qy        422 NKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPIC 481
              |||  |||||  |||  ||||  |||  |||  |||  |||  |||  |||  |||  |||
Db        300 NKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPIC 359

Qy        482 VGLGLYAQWGSGKSFLLKKLEDEMKTFAGQQTEPLPQFSWLIVFLTLLLVGGLGLVFAFPV 541
              |||||||||||||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        360 VGLGLYAQWGSGKSFLLKKLEDEMKTFAGQQTEPLPQFSWLIVFLTLLLVGGLGLVFAFPV 419

Qy        542 DTNLAIAISLSLFLALIYIFFIVYF--R---WVWW--V---VPPLVPLF 581
              |||  |||  |||  |||  |||  |  |   |||  |   |||  |||
Db        420 HFNLGIAVSLSLFLALIYIFFIVYF--R---WVWW--V---VPPLVPLF 459

Qy        582 NPPELPEQTTKALPVRFLFTDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFR 641
              |||||  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        460 NPPELPEQTTKALPVRFLFTDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFR 519

Qy        642 RTEESQGKKKWK 653
              |||  |||  |||
Db        520 KTEDTQGKKKKK 531


RESULT 13
US-10-[illegible]-571-8
; Sequence 8, Application US/10021[illegible]71
; Publication No. US20030166056A1
; GENERAL INFORMATION:
;  APPLICANT: CHAO, Moses V.
;  APPLICANT:  KONG, Haeyoung
;  TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
;  TITLE OF INVENTION:  RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI
;  FILE REFERENCE: CHAO11A
;  CURRENT APPLICATION NUMBER: US/10/021,571
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/256,909
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;    LENGTH: 1498
;    TYPE: PRT
;    ORGANISM: Drosophila
US-10-021-571-8

  Query Match             26.6%;  Score 2376.5;  DB 12;  Length 1498;
  Best Local Similarity   37.6%;  Pred. No. 7.1e-173;
  Matches  562;  Conservative  272;  Mismatches  527;  Indels  117;  Gaps   36;

Qy          1 MSVLISQSVINYVEEENIPALKALLE-KCKDVDERNECGQTPLMLAAEQGNVEIVKELLK 59
              |  |  |:::| |::  :|  |:|:|: :  :::::|   | ||: | :|    :|  |
Db         23 MGSLGHSALLQYIDNNDISGLRAILDSRHLTIDDRDENATTVLMVVAGRGLTAFVREFLA 82

Qy         60 NGANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDV 119
              ||:   ||  ||||||: ||: :|: :|: ||  :| :|||||||||: ||| |:: |::
Db         83 RGADVQAEDLDNWTALLCASRNGHLDVVQLLLDHGAEVEHRDMGGWTSLMWAAYRGHTEL 142

Qy        120 VELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAA 179
              ||:  ||: :  |  : ::|||':|: |.' ||:| |||||  |||||: | ||
Db        143 VRLLLDKGADGNAHG-NYHLGALIWAAGRGYKDIVELLVQRGAKVNVGDKYGTTALVWAC 201

Qy        180 RKGHLEVVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNT 239
              :|:|:  |  ||:  |  |  | :  |  |  | :  |  ,|:: |  ||:
Db        202 RRGNVEIVDTLLKAGANVDTAGMYSWTPLLVAAAGGHTDCVSSILEKKPNVNALDKDGMT 261

Qy        240 ALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIR 299
              |' |||:|  :  |: | |: |  || | :| || |: |||:|:||:|||:
Db        262 ALCIASREGFQDIAASLIAAGAYINIQDRGADTPLIHAVKAGHRTVVEALLKKHADVDIQ 321

Qy        300 GQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKG 359
              |:  |||:| ||:|:  :|: :|  ||| |  ||||:|:|:|: : ||:|:, |'':
Db        322 GKDRKTAIYTAVEKGHTPIVKILLATNPDLESATKDGDTPLLRAVRNRNLEIVHLLLDRK 381

Qy        360 AKVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSI 419
              |||:  ||:|'| |:|:|: '':  : . ||||:.  :|||| ||||||:'|||  ':|:|
Db        382 AKVTASDKRGDTCLHIAMRARSKTIVEALLRNPKHSQLLYRANKAGETPYNIDSLHQKTI 441

Qy        420 LTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKK 479
              '  |:|| |:  |    ||.|:|||||||:|| ||:  || |||||:|||||| |
Db        442 LGQVFGARRLNTNEDSEGMLGYELYSSALADVLSEPTLTTPITVGLYAKWGSGKSFLLNK 501

Qy        480 LEDEMKTFAGQQTEPLPQFSWLIVFLTLLLVGGLGLVFAFPV-DTNLAIAISLSFLALIY 538
              ' |||  |:    ||  : | |:  : |:  :  :                 : :: :: || . '
Db        502 LSDEMNNFAPQWAEPPIRTSGLLPIVGLHVALLIGTIVGLSTWSAVVGVSAAVGFLLLAY 561

Qy        539 IFFIVIYFGGRREGESWNWAWALSTRLARHIGYDELLFKLMFVNPPELPEQTTKALPVRF 598
              :  :: :  :  ||:::  | : :  , :: : :||  |:  ::| ||||
Db        562 LLLAAVRYCNYQ--MDMQWAYSVQHGLEKRMTRLRLILQVAFCHPPG-PQSDSQAKPVRF 618

Qy        599 LPTDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFRTE--ESQGKKKWKKTC 656
              : | |  |: :: |:| | || | :|:||:|| || :  :|:: |
Db        619 HFAEANSASPT-GDGAVAHMLAALLDAIESHYGWLATRLYRAFRPKCLKVDVGWRWRRMC 677

Qy        657 CLPSFVIFLFIVGCIIAGITLLAIF-----RVDPKHLTVNAILISIASVVGLAFVLNCRT 711
              |:|  :||   :  :: ||:|  :       : :|: |  | ||:|:|     :
Db        678 CIPIVLIFELALVTVVTGISLTVAYPTFADEKEKEHILV--ALYVIAAVMGTLICTHLHV 735

Qy        712 WWQVLDSLLNSQRKRLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLV 771
              :|  ||  |  : :|  |  :  : :|  || :|  | | :|:|| |:|||
Db        736 LAKVFVSLFTSHIRVLKRAV----RSSESAPLTMLGAEVAVMTDMVKCLDAFTNQQSRLV 791

Qy        772 VIIDGLDACEQDKVLQMLDTVRVLFS--KGPFIAIFASDPHIIIKAINQNLNSVLRDSNI 829
              :|| ||:  :|:| :|: |: | |   ||: : : |||:| ||  |  :  : :
Db        792 GVIDALDSCDTERILTLLNAVQTLLSSPNRPFVLLISVDPHVIAKAAEANSRRLFTEGGI 851

Qy        830 NGHDYMRNIVHLPVFLNSRGLSNARKFLVTS-----ATNGDITCSD--TTGTQEDTDRRV 882
              |||:||:-'|||:| : |'  ::  :|:  :  ||  |  | |  : ||:
Db        852 GGHDFLRNLVHLPVYLQNSGLRKVQRAQMTALLFKRSGGGDYQTDDGPTLG-HSVSARRL 910

Qy        883 SQNS--LGEMTKL--------GSKTALNR-------RDTYRRRQMQRTITRQMSFDLTKL 925
              !  |  :  :|         |  |:          : :|  |  :|:      | :::
```

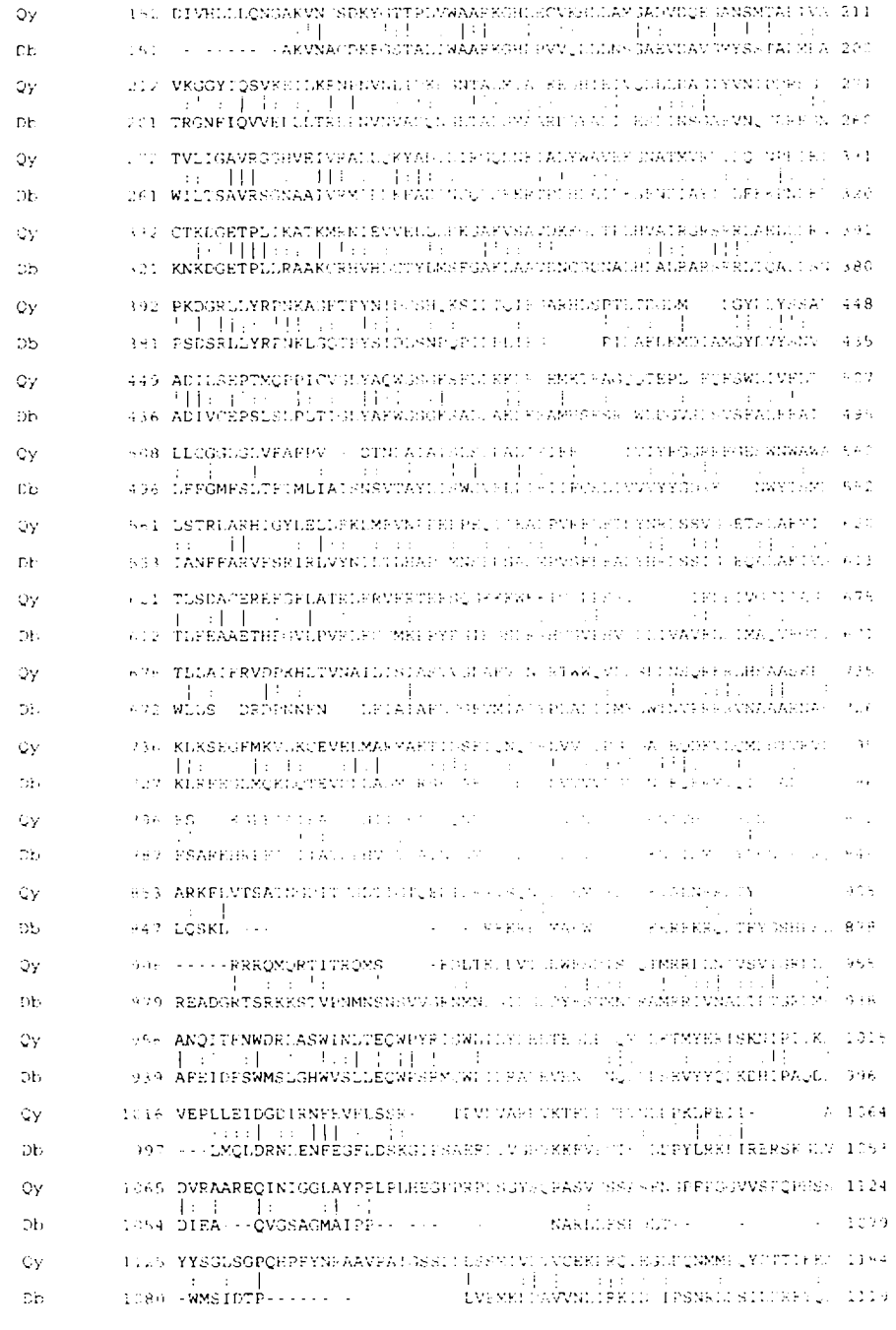
RESULT 14
US-10-143-819-9
; Sequence 9, Application [illegible]
; Publication No. US2003-0044-[illegible]
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX-ASSOCIATED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/143,819
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/171,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. [illegible]
US-10-143-819-9

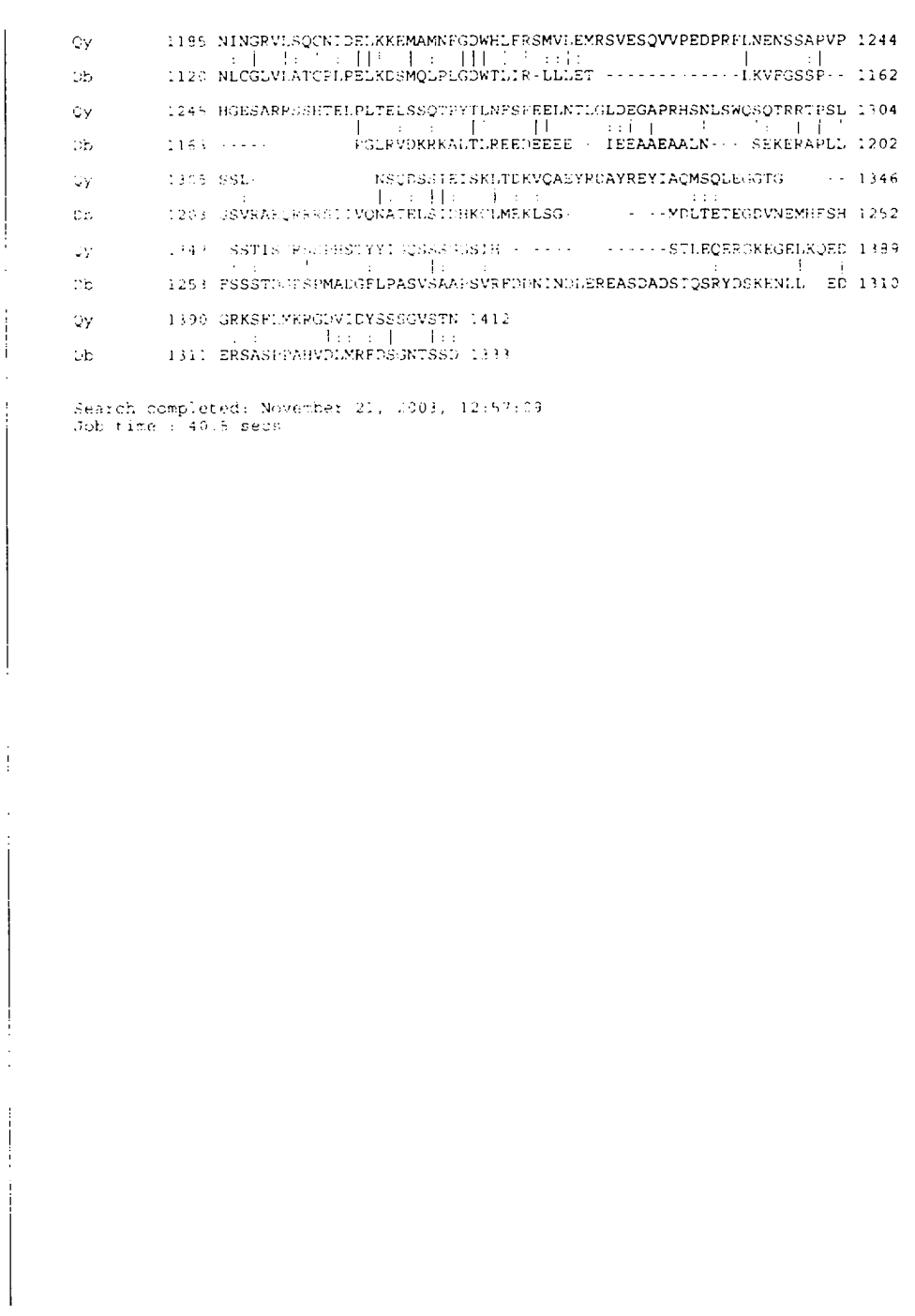
Query Match 25.5%; Score [illegible]; DB 1; Length [illegible];

Best Local Similarity 86.6%; Pred. No. 9.9e-165;
Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;

RESULT 15
PR [illegible]
; Sequence 6, Application US/10/021,571
; Publication No. US2002-0164[illegible]
; GENERAL INFORMATION:
; APPLICANT: MEAD, Steven V.
; APPLICANT: KONG, Haeyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN /
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL AN
; FILE REFERENCE: CHA01A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1398
; TYPE: PRT
; ORGANISM: Caenorhabditis Elegans
US-10-021-571-6

Query Match 20.9%; Score 1848.5; DB 12; Length 1398;
Best Local Similarity 31.5%; Pred. No. 2.5e-132;
Matches 461; Conservative 272; Mismatches 459; Indels 271; Gaps 37;





Search completed: November 21, 2003, 12:57:09
Job time : 40.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 12:11:54 ; Search time 44.5 Seconds
(without alignments)
6117.215 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....PIHAASSESTGFSEERRSIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 1-8726571 residues

Total number of bits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

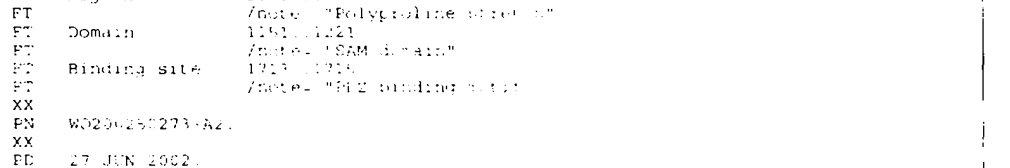
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8853	100.0	1715	23	AAE25143	Rat ARMS protein.
2	8839.5	99.8	1762	23	AAU96841	Rat kidins220 prot
3	8526	96.3	1763	23	AAU80244	Rat kidins220 prot
4	8266	93.4	1715	22	AAM38993	Human polypeptide
5	8266	93.4	1715	22	AAM39025	Human polypeptide
6	8266	93.4	1715	23	AAE25144	Human ARMS protein
7	8266	93.4	1715	23	AAU96840	Human kidins220 pr
8	8266	93.4	1715	24	AAE32126	Human cytoskeleton
9	2961.5	33.5	705	22	AAG75604	Human colon cancer
10	2782.5	31.4	551	22	AAE01035	Human death domain
11	2408	27.2	543	22	AAB95191	Human protein sequ
12	2376.5	26.8	1498	22	ABB64857	Drosophila melanog
13	2376.5	26.8	1498	23	AAE25146	Fruit fly ARMS pro
14	2260	25.5	513	22	AAE03645	Human extracellula
15	2150	24.3	624	23	ABG08697	Novel human diagno
16	1648.5	20.9	1398	23	AAE25145	Worm ARMS protein.
17	614	6.9	120	22	AAM40811	Human polypeptide
18	599.5	6.8	129	22	AAB94785	Human protein sequ
19	549	6.2	109	21	AAB56277	Human secreted pro
20	538.5	6.1	1724	23	AAO20513	Protein of APP rel
21	534.5	6.0	1872	22	AAM79160	Human protein SEQ
22	531.5	6.0	986	22	ABG31785	Human ankyrin 4 po
23	531.5	6.0	1250	24	AAE33667	Human structural a
24	526	5.9	4274	22	ABG00972	Novel human diagno
25	521	5.9	4386	22	ABG07375	Novel human diagno
26	521	5.9	4397	22	ABG21944	Novel human diagno
27	512	5.8	1536	24	ABU11523	Human MDDT polypep
28	511	5.8	456	22	ABU53033	Cell structure and
29	504	5.7	2443	22	ABB60521	Drosophila melanog
30	500.5	5.7	187	22	AAM40779	Human polypeptide
31	492.5	5.6	784	22	ABB62206	Drosophila melanog
32	475.5	5.4	489	22	ABU53032	Cell structure and
33	471	5.3	447	22	ABU53031	Cell structure and
34	471	5.3	2119	22	ABB64823	Drosophila melanog
35	467.5	5.3	1549	22	ABB58328	Drosophila melanog
36	467.5	5.3	1549	22	ABB67412	Drosophila melanog
37	466.5	5.3	1745	19	AAW70608	Full length ankyri
38	466.5	5.3	1745	19	AAW76776	D. immitis ankyrin
39	466.5	5.3	1745	21	AAB11589	D. immitis ankyrin
40	466.5	5.3	1745	23	AAO21368	Ankyrin protein se
41	459	5.2	743	23	ABB97453	Novel human protei
42	451.5	5.1	447	22	ABU53030	Cell structure and
43	450	5.1	1133	22	ABG08071	Novel human diagno
44	449.5	5.1	1088	22	AAB62429	Human kidney ankyr
45	449.5	5.1	1094	24	AAB33688	Human structural a

ALIGNMENTS

RESULT 1
AAE25143
ID AAE25143 standard; Protein; 1715 AA.
XX
AC AAE25143;
XX
DT 30-OCT-2002 (first entry)
XX
DE Rat ARMS protein.
XX
KW Ankyrin repeat rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW diagnostic; imaging; rat.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT Region 1..390
FT /note= "Ankyrin repeat"
FT Region 354..493
FT /note= "N-terminal region"
FT Domain 496..519
FT /note= "Transmembrane domain"
FT Domain 525..546
FT /note= "Transmembrane domain"
FT Region 573..636
FT /note= "Loop between TM2 and TM3"
FT Domain 661..680
FT /note= "Transmembrane domain"
FT Domain 698..710
FT /note= "Transmembrane domain"
FT Region 747..854

```
FT                   /note= "Carboxy tail"
FT   Region          940..1040
FT                   /note= "Carboxy tail"
FT   Region          1081..1126
FT                   /note= "Polyproline stretch"
FT   Domain          1151..1221
FT                   /note= "SAM domain"
FT   Binding site    1715..1715
FT                   /note= "PKC binding site"
XX
PN   WO200250273-A2.
XX
PD   27-JUN-2002.
XX
PF   19-DEC-2001; 2001WO-US49843.
XX
PR   21-DEC-2000; 2000US-0258298P
XX
PA   (UYNY ) UNIV NEW YORK STATE.
XX
PI   Chao MV, Kong H;
XX
DR   WPI; 2002-508900/54.
DR   N-PSDB; AAD41035.
XX
PT   Isolated ankyrin repeat rich membrane spanning (ARMS) polypeptide that
PT   is a target for phosphorylation by neurotrophin and ephrin receptor
PT   tyrosine kinases, useful as a marker for growth cones
XX
PS   Claim 1; Page 87-94; 133pp; English.
XX
CC   The invention relates to ankyrin repeat rich membrane spanning (ARMS
CC   protein which is a down stream target of neurotrophin and ephrin
CC   receptor tyrosine kinases. ARMS DNA is used for visualising the
CC   growth cone of neurons. ARMS protein is useful as an indicator of
CC   the biological activity of neurotrophins and ephrins, as a marker
CC   for neuronal cells which have the ability to undergo continued
CC   synaptic changes through adult life or for the presence and
CC   distribution of ARMS in such neuronal cells, or as a marker for
CC   growth cones. ARMS DNA is useful in diagnostic and imaging methods.
CC   The present sequence is rat ARMS protein.
XX
SQ   Sequence  1715 AA;
```



```
Qy          1 MSVLISQSVINYVEENIPALAALLEK
Db          1 MSVLISQSVINYVEENIPALAALLEK 60

Qy         61 GANCNLEDLDNWTALISASKEHIHIVEELIS
Db         61 GANCNLEDLDNWTALISASKEHIHIVEELIS 120

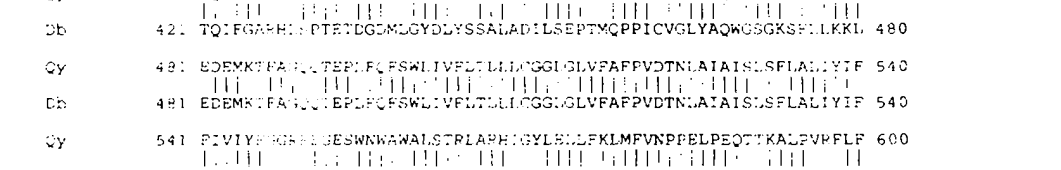
Qy        121 ELLLSHGANPSVTGLQYSVYEIIWAAGH
Db        121 ELLLSHGANPSVTGLQYSVYEIIWAAGH 180

Qy        181 KGHLECVKHLLAMGADVDQEGANSMTA
Db        181 KGHLECVKHLLAMGADVDQEGANSMTA 240

Qy        241 LMIASKEGHIEIVQDLLDAGTYVNIPDR
Db        241 LMIASKEGHIEIVQDLLDAGTYVNIPDR 300

Qy        301 QDNKTALYWAVEKGNATMVRDLLQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
Db        301 QDNKTALYWAVEKGNATMVRDLLQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy        361 KVSAVDKKGDTPLHVAIRGRSPRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
```



```
Db        361 KVSAVLKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy        421 TQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGGKSFLLKKL 480
Db        421 TQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGGKSFLLKKL 480

Qy        481 EDEMKTPAFQLTEPLQFSWLIVFLTLLLVGGIGLVFAFPVDTNLAIAISLSFLALIYIF 540
Db        481 EDEMKIPAFQLTEPLQFSWLIVFLTLLLVGGIGLVFAFPVDTNLAIAISLSFLALIYIF 540

Qy        541 PIVIYFQGFFLGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 600
Db        541 PIVIYFQGFFLGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 600

Qy        601 TDYNPLSSVSGETSLAEMIATLSPACERFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 660
Db        601 TDYNPLSSVSGETSLAEMIATLSPACERFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 660

Qy        661 PVIFLFIVSCIIAGITLLAIPRVDPKHLTVNAILISIASVVGLAFVLNCFTWWQVLDSLL 720
Db        661 PVIFLFIVSCIIAGITLLAIPRVDPKHLTVNAILISIASVVGLAFVLNCFTWWQVLDSLL 720

Qy        721 NSQRKSLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIDGLDAC 780
Db        721 NSQRKSLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIDGLDAC 780

Qy        781 EQDKVLQMLGTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Db        781 EQDKVLQMLGTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Qy        841 LPVFLNSPKLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900
Db        841 LPVFLNSPKLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900

Qy        901 RRDTYPRPLMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960
Db        901 RRDTYPRPLMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960

Qy        961 PNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db        961 PNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020

Qy       1021 RIDGQTPNIAVPLSSPRIPVLVAPSVKTPLPTTVNLLPKLREIIADVRAAREQINIGGLAY 1080
Db       1021 RIDGQTPNIAVPLSSPRIPVLVAPSVKTPLPTTVNLLPKLREIIADVRAAREQINIGGLAY 1080

Qy       1081 PELPLPLPLPPPSGYSQPASVCSSASFNGPFPGVVSPQPHSSYYSGLSGPQHPFYNRA 1140
Db       1081 PELPLPLPLPPPSGYSQPASVCSSASFNGPFPGVVSPQPHSSYYSGLSGPQHPFYNRA 1140

Qy       1141 AVPATGSSLLLSSMTVDVVCEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELK 1200
Db       1141 AVPATGSSLLLSSMTVDVVCEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELK 1200

Qy       1201 KEMAMNPGIWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPLT 1260
Db       1201 KEMAMNPGIWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPLT 1260

Qy       1261 ELSSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLT 1320
Db       1261 ELSSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLT 1320

Qy       1321 DKVQAEYPDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERG 1380
Db       1321 DKVQAEYPDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERG 1380

Qy       1381 KEGELKQEDGRKSPLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKS 1440
Db       1381 KEGELKQEDGRKSPLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKS 1440

Qy       1441 SERPSLPGTDLKLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIRTKRLKAKQRECASP 1500
```

```
Db        1441 SERPSLFQTDLKLKGGGLEYQKLPSDEIASGTGRVQITPHCSKMIRTKRLKAKQRETAHF 1500

Qy        1561 QEHSAEPIRTFIKAKEYLSDALLDKKDCDPSDVFSNESSPNHSLHNEAAIDSQLEKANLI 1620

Db        1501 QEHSAEPIRTFIKAKEYLSDALLDKKDCDPSDVFSNESSPNHSLHNEAAIDSQLEKANLI 1560

Qy        1621 LNRTPSTVTLNNNTAPTNRANQNFDEIEYIRETSQVIIRPGFNFNPTAVQNENLKSMADK 1680

Db        1561 ELEDEGHSGKRGMPHSLSGLQPTIARMSISEKKSFSECSLIASSPERSWFACQFAYN 1620

Qy        1681 RSQRSSYTRLSKDASELHAASNESTGFGEERENIL 1715

Db        1621 LNRTPSTVTLNNNTAPTNRANQNFDEIEYIRETSQVIIRPGFNFNPTAVQNENLKSMADK 1680

Db        1681 RSQRSSYTRLSKDASELHAASNESTGFGEERENIL 1715
```

RESULT 2
AAU96841
ID AAU96841 standard; Protein; 1762 AA.
XX
AC AAU96841;
XX
DT 10-JUL-2002 (first entry)
XX
DE Rat kidins220 protein from Genbank AAG35185.
XX
KW Kidins220; kinase D interacting substrate of 220kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX
OS Rattus sp.
XX
PN WO200220786-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000US-23844P.
XX
PA (IMCR) IMPERIAL CANCER RES TECH LTD.
XX
PI Schiavo G, Illeyton T;
XX
DR WPI; 2002-371879/43.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying agent useful in treating
PT cancer and neurodegenerative diseases
XX
PS Disclosure; Fig 18; 228pp; English.
XX
CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by GenBank Accession No. AF293796. Also included are
CC a polynucleotide encoding KIdins220 provided that polynucleotide is
CC not any one of the clones corresponding to the 41 GenBank Accession Nos.
CC given in the specification; an expression/replicable vector comprising
CC the polynucleotide; a recombinant host cell containing the polynucleotide
CC or vector; an anti-Kidins220 antibody used in the preparation of
CC Kidins220); an agent (A1) which modulates activity of Kidins220,
CC protein kinase D (PKD) or interaction between PKD and Kidins220;
CC a mutant animal transgenic for kidins220; the use of an agent capable of
CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
CC reagent for diagnosing or prognosing cancer or for monitoring the
CC progression of cancer in a patient; a compound comprising a moiety that
CC selectively binds to Kidins220 or its variant and another moiety; and a
CC nucleic acid molecule encoding the compound. The antibody is useful for
CC modulating the activity of Kidins220, where the antibody prevents a
CC region of Kidins220 interacting from another protein, or reduces the
CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
CC kidins220 polynucleotide is useful for identifying an agent which
CC modulates the promoter activity of the polynucleotide, and Kidins220 is
CC useful for identifying an agent which modulates activity of Kidins220,
CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
CC therapy or for diagnosing or monitoring progression of cancer in a
CC patient. The agent is useful in manufacture of medicament, for use in
CC treatment of neurodegenerative disease, and the agent which inhibits
CC function of Kidins220 gene or its product is useful for treating cancer
CC which is glioblastoma multiforme or prostate cancer. The compound is
CC useful for imaging cancer in an individual, and for diagnosing or
CC prognosing, and also for treating an individual. The present
CC sequence represents Rat kidins220 protein from Genbank accession number
CC AAG35185.
XX
SQ Sequence 1762 AA;

Query Match 97.6%; Score 8639.5; DB 23; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

```
Qy          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60
              |  |||  ||||. |||,|||||||||||||||||||||||||||||||||||||||||||
Db          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60

Qy         61 GANCNLEDKLNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
              |||||||| ||| ||||| |||||||||||| |||||||||||||||||||||||||||||
Db         61 GANCNLEDKLNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
              |||| ||| |||| ||||||||||| |||| |||||| |||| |||| |||| | ||||
Db        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 179

Qy        181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
              |||||||||  ||| |||| ||||||||||||||| ||||||||||||| |||||||||
Db        180 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 239

Qy        241 LMIASKEGHIEIVQLLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
              ||| | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 LMIASKEGHIEIVQLLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 299

Qy        301 QDNKTALWWAVERGNATMVRILQQNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
              || |||  |||  ||||  ||| |||| |||| ||||| |||| |||| | || |||
Db        300 QDNKTALWWAVERGNAIMVRILQQNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 359

Qy        361 KVSAVDKKGRTPLHVAIRGRSPRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
              |||| ||  |||| ||| |||| |||| |||||||||||||||||||||||||||||||
Db        360 KVSAVDKKGRTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 419

Qy        421 TQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480
              ||| |||| |||| |||| |||||||||||||||||||||||||||||||||||||||||
Db        420 TQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 479

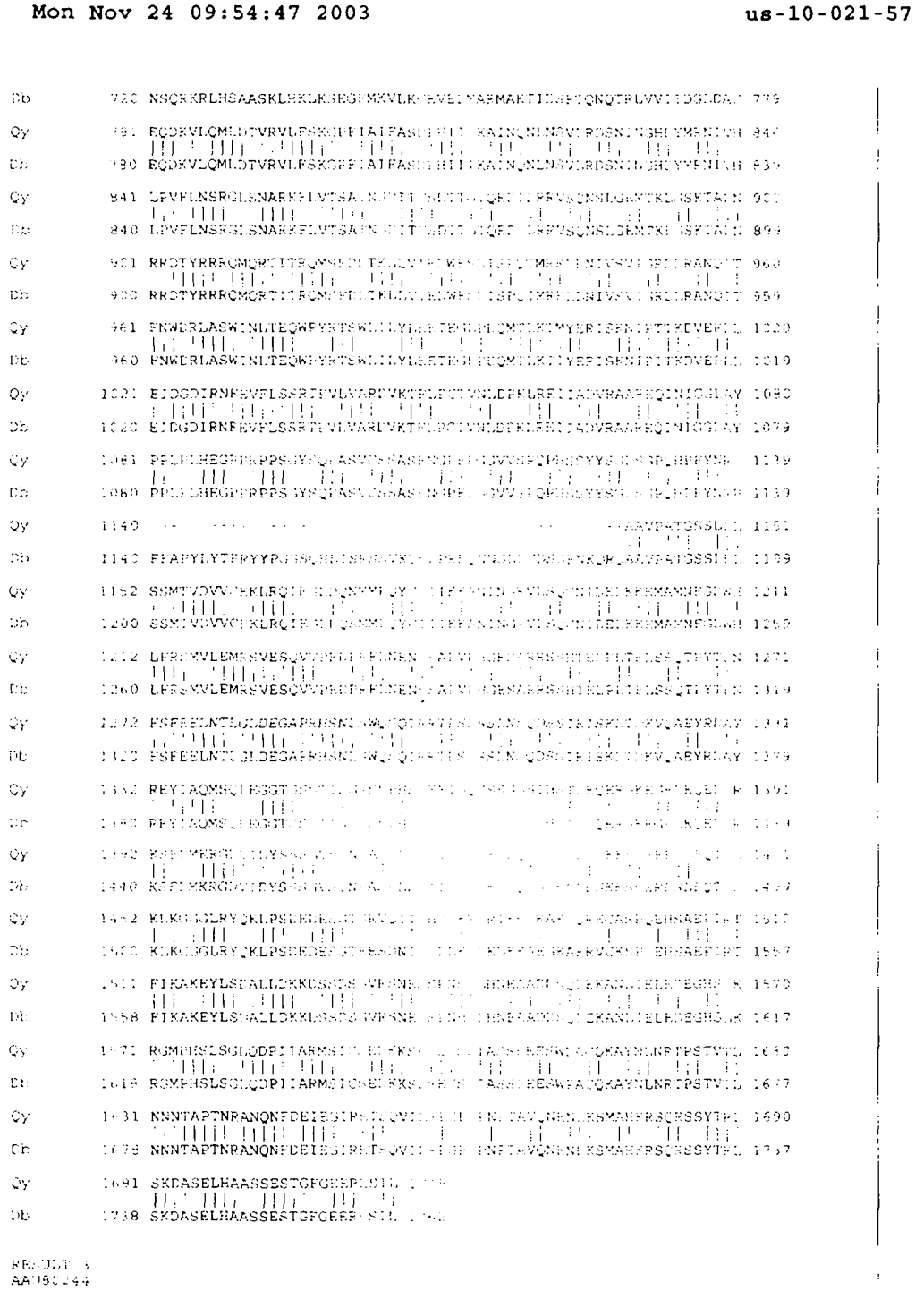
Qy        481 EDEMKTFAIQQTEPLFQFSWLIVFLTLLLCGGLGLVFAFPVDTNLAIAISLSFLALIYIF 540
              ||| |||  |||| ||| | ||| |||| |||| |||||||||||||||||||||||||
Db        480 EDEMKTFAIQQTEPLFQFSWLIVFLTLLLCGGLGLVFAFTVDTNLAIAISLSFLALIYIF 539

Qy        541 FIVIYFGGRFEGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 FIVIYFGGRFEGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 599

Qy        601 TDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 660
              |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 659

Qy        661 FVIFLFIVICIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720
              ||| ||| ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        660 FVIFLFIVICIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 719

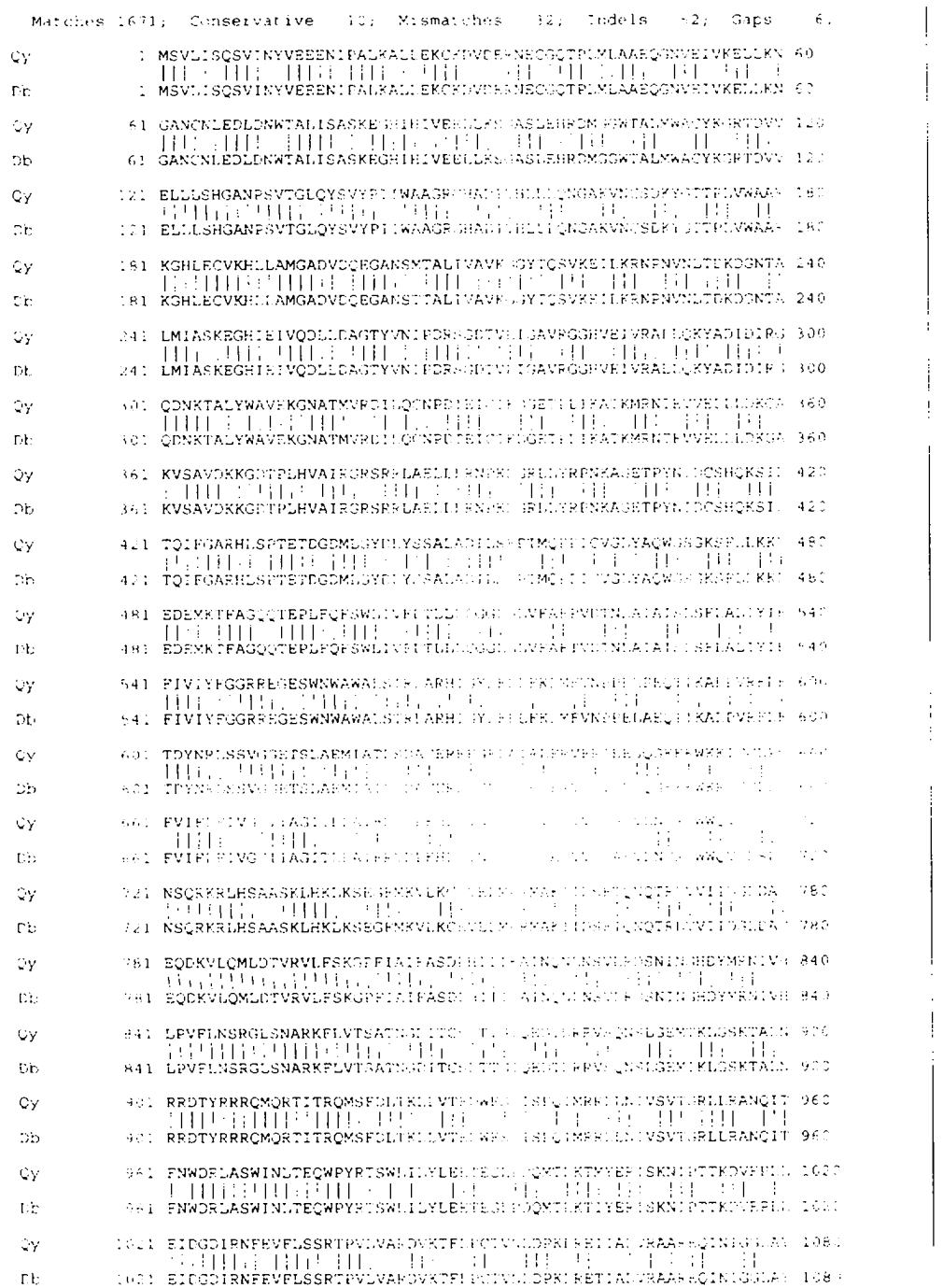
Qy        721 NSQRKPLH AASKLHKHKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780
              ||| |||| ||| |||| |||||||||||||| |||||||||| ||||||||||||||
```



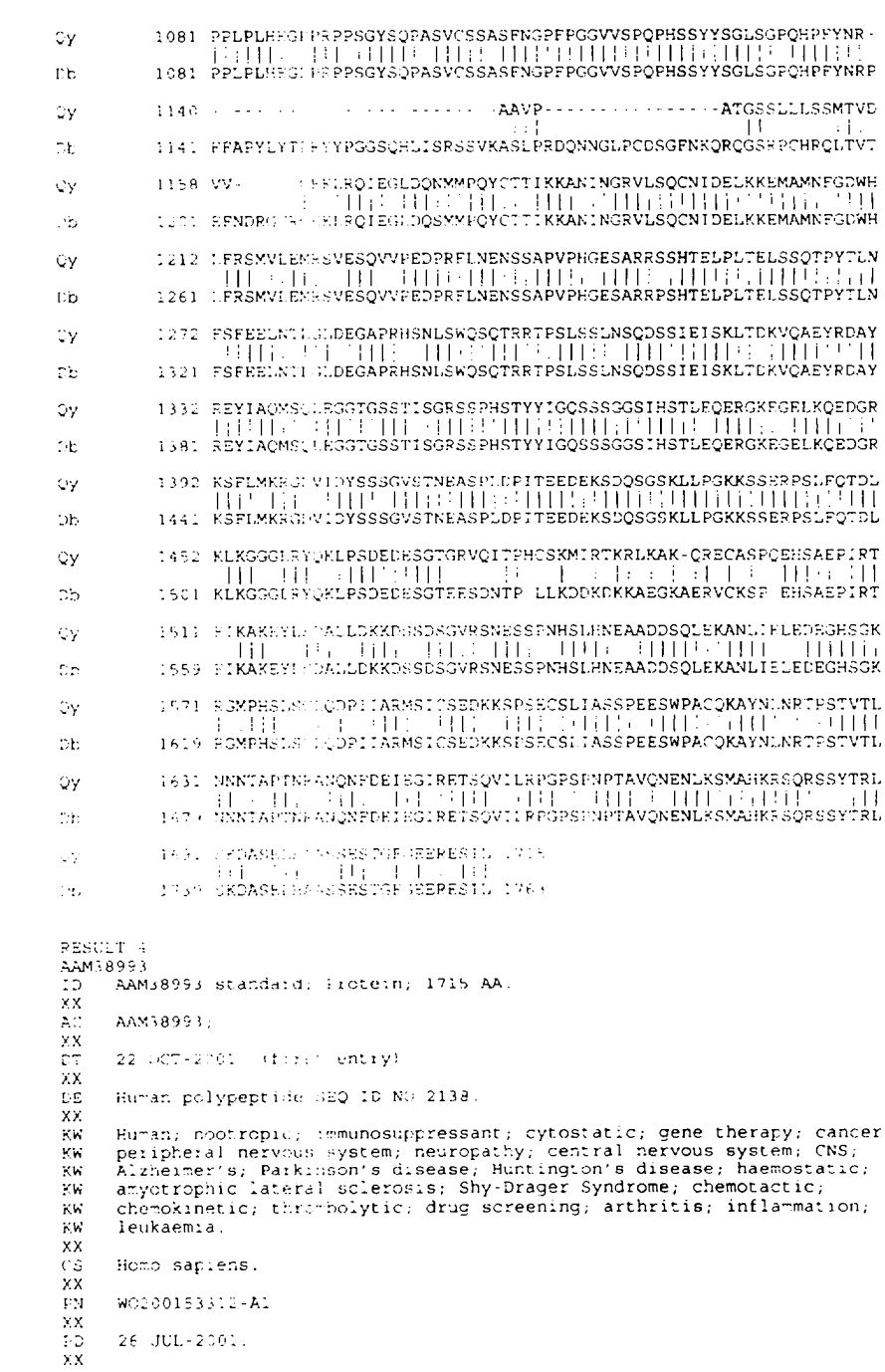
RESULT 5
AAU80244

ID AAU80244 standard; Protein; 1763 AA.
XX
AC AAU80244;
XX
DT 30-JUL-2002 (first entry)
XX
DE Rat kidins220 protein.
XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX
OS Rattus sp.
XX
PN WO200207786-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 04-SEP-2000; 2000US-230449P.
XX
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Sanchez C, Iglesias T;
XX
DR WPI; 2002-371879/40.
DR N-PSDB; AAK51214.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating
PT cancer and neurodegenerative diseases -
XX
PS Claim 33; Fig 1; 108pp; English.
XX
CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by GenBank Accession No. AB033076. Also included are
CC a polynucleotide encoding Kidins220 provided that polynucleotide is
CC not any one of the clones corresponding to the 61 GenBank Accession Nos.
CC given in the specification; an expression/replicable vector comprising
CC the polynucleotide; a recombinant host cell containing the polynucleotide
CC or vector; an anti-Kidins220 antibody used in the preparation of
CC Kidins220; an agent (A1) which modulates activity of Kidins220,
CC protein kinase D (PKD) or interaction between PKD and Kidins220;
CC a transgenic animal, transgenic for Kidins220, the use of an agent capable of
CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
CC reagent for diagnosing or prognosing cancer or for monitoring the
CC progression of cancer in a patient; a compound comprising a moiety that
CC selectively binds to Kidins220 or its variant and another moiety, and a
CC nucleic acid molecule encoding the compound. The antibody is useful for
CC modulating the activity of Kidins220, where the antibody prevents a
CC region of Kidins220 interacting from another protein, or reduces the
CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
CC kidins220 polynucleotide is useful for identifying an agent which
CC modulates the promoter activity of the polynucleotide, and Kidins220 is
CC useful for identifying an agent which modulates activity of Kidins220,
CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
CC therapy or for diagnosing or monitoring progression of cancer in a
CC patient. The agent is useful in manufacture of medicament, for use in
CC treatment of neurodegenerative disease, and the agent which inhibits
CC function of Kidins220 gene or its product is useful for treating cancer
CC which is glioblastoma multiforme or prostate cancer. The compound is
CC useful for imaging cancer in an individual, and for diagnosing or
CC prognosing, and also for treating an individual. The present
CC sequence represents Rat kidins220.
XX
SQ Sequence 1763 AA;

Query Match 96.1%; Score 8526; DB 23; Length 1763;
Best Local Similarity 94.7%; Pred. No. 0;



Matches 1671; Conservative 10; Mismatches 12; Indels 12; Gaps 6.



RESULT 4
AAM38993
ID AAM38993 standard; protein; 1715 AA.
XX
AC AAM38993;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2138.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1
XX
PD 26-JUL-2001.
XX

PF 26 DEC 2000; 2000WO-US34263.
XX
PR 21 JAN-2000; 2000US-0488725.
PR 25 APR 2000; 2000US-0552317.
PR 04-JUL-2000; 2000US-0598043.
PR 19 JUL 2000; 2000US-0620812.
PR 03-AUG-2000; 2000US-0635455.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT 2000; 2000US-0693036.
PR 29 NOV 2000; 2000US-0727144.
XX
PA (HYSE) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Xu S, Qian XB, Ren F, Wang F,
PI Wang C, Wang Z, Wehrman T, Xie Y, Xue AY, Yang Y, Zhang Z,
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT,
XX
DR WPI; 2001-442053/47.
DR N-PSDB; AAI58149.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries.
XX
PS Example 4; SEQ ID NO 2134; 1015pp; English.
XX
CC The invention relates to human secreted proteins (AAI57038-AAI57 [illegible]) and
CC the encoded polypeptides (AAM39444-AAM42[illegible] with neurotrophic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous system injury, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy Drager Syndrome. Other uses include the
CC utilisation of the activities such as immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC CNS disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1715 AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible]; Mismatches [illegible]; Conservative [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 MSVLISQSVINYVEEENIPALKALDEK FVEKNSNENIELIFKVIEKLPKY 60
Db 1 MSVLISQSVINYVEEENIPALKALDEK FVEKNSNENIELIFKVIEKLPKY 60

Qy 61 GANCNLEDLDNWTALISASKEGHIEIVEFILKHASLEHPKDSWTALMWACYKGRTDVV 120
Db 61 GANCNLEDLDNWTALISASKEGHVEIVEFLLKHANLEHPKDSWTALMWACYKGRTDVV 120

Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAGKHNAIVRLLIQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAGKHNAIVRLLIQNGAKVNCSDKYGTTPLVWAAR 180

Qy 181 KGHLECVKHLLAMGADVDQRGANSMTALIVAVF YTQSVKEILKRNPNVNLTLKDGNTA 240
Db 181 KGHLECVKHLLAMGADVDQRGANSMTALIVAVF YTQSVKEILKRNPNVNLTLKDGNTA 240

Qy 241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGQTVLIGAVRSGHVEIVRALLQKYADIDIRS 300
Db 241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGQTVLIGAVRSGHVEIVRALLQKYADIDIRS 300

Qy 301 QDNKTALYWAVEKGNATMVRDILQNRPDTEICLKDGETPLIKATKMPNIEVVELLLKKGA 360
Db 301 QDNKTALYWAVEKGNATMVRDILQNRPDTEICLKDGETPLIKATKMRNIEVVELLLKKGA 360

Qy 361 KVSAVDKKGDTPLHIAIRGRSRRLAELLLRNPKLGRLLYRENKAGETPYNIDGSHQKSIL 420
Db 361 KVSAVDRKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDGSHQKSIL 420

Qy 421 TQIPGAKHSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480
Db 421 TQIPGARHSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480

Qy 481 EDEMKTRAIQTIEPLFQFSWLIVFDTLLLCGGLGLVFAFPVDTNLAIAISLSFLALIYIF 540
Db 481 EDEMKTRAIQTIEPLFQFSWLIVFDTLLLCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540

Qy 541 FIVIYFGSWNWAWAISTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVREFLF 600
Db 541 FIVIYFGSWNWAWVLSTRLARHIGYLELLLKLMFVNPPELPEQTTKALPVREFLF 600

Qy 601 TDYNRLASVGETSLAEMIATLSDACEREPGFLATRLFRVFRTEFSQGKKKWKKTCCLPS 660
Db 601 TDYNRLASVGETSLAEMIATLSDACEREPGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660

Qy 661 FVIFLFIVFIIAGITILAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720
Db 661 FVIFLFIIFIISGITILAIFRVDPKHLCVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy 721 NSQRKRIHDAASKLHKLKSEGFMKVLKCEVELMARVAKTIDSFTQNQTRLVVIIDGLDAC 780
Db 721 NSQRKRIHDAASKLHKLKSEGFMKVLKCEVELMARVAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy 781 EQDKVLQVICTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Db 781 EQDKVLQVICTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Qy 841 LPVFLNSRKSNARKFLVTSATNGDVPCSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRKSNARKFLVTSATNGDVPCSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900

Qy 901 KRDTYPRKIMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960
Db 901 PRDTYPRKIMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960

Qy 961 PNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 961 PNWDRLASWINLTEQWPYRTSWLILYLEETEGIPDQMTLKTIYERISKNIPTTKDVEPLL 1020

Qy 1021 EIDGDIENEVFLSSRTPVLVARDVKTFLPCTVNLDPKLREIIADVRAAREQISIGGLAY 1080
Db 1021 KIDGDIENEVFLSSRTPVLVARDVKVFLPCTVNLDPKLREIIADVRAAREQISIGGLAY 1080

Qy 1081 PPLPLHEGPRAPSGYSQPPSVCSSTSFNGPFAGGVVSPQPHSSYYSGLSGPQHPFYNRA 1140
Db 1081 PPLPLHEGPRAPSGYSQPPSVCSSTSFNGPFAGGVVSPQPHSSYYSGMTGPQHPFYNRG 1140

Qy 1141 AVPATGSSLLSSMTVFVVCEKLPQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELK 1200
Db 1141 SGPAPGEVVLLNSLNVDAVCEKLKQIEGLDQSMLPQYCTTIKKANINGRVLAQCNIDELK 1200

Qy 1201 KEMAMNPGIWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHT 1260
Db 1201 KEMNMNPGIWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHT 1260

Qy 1261 ELSSQTPYILNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLT 1320
Db 1261 ELSSQTPYILNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLT 1320

Qy 1321 DKVQAEYEIAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERG 1380
Db 1321 DKVQAEYEIAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKG 1380

Qy 1381 KEGELKQEIGRKSFLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKS 1440
Db 1381 KDSEPKFLIGRKSFLMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKS 1440

Qy 1441 SERPSLFQIFKLKGGSLRYQKLFSDEDESGTGRVQITPHCSKMIRTKRLKAKQRECASP 1500

```
Db       1441 SERSSLFQTDLKLKGSGLRYQKLPSDEDESGTEESLNTP-LLKDDKLRKAEGKVERVPKS 1499

Qy       1501 QERSAEPIRTFIKAKEYLSDALLDKKDSSDSTVRSNESSPNHSLHNEAADDSQLEKANLI 1560
              ||: :|||| .:.|||||: |||: ||| ||: ||| ||| ||| :||| ||
Db       1500 PERSAEPIRTFIKAKEYLSDALLDKKDSSDSGVPSESESSPNHSLHNEVADDSQLEKANLI 1559

Qy       1561 ELEDEGHSGKRGMPHSLSSLQDPIIARMSIGSLRKSPSSCSLIASSPEESWPACQKAYN 1620
              .||: :|||||:: |||| . ||| :.||| ||| ||| ||| :|:.||| ||
Db       1560 ELEDDSHSGKRGIPHSLSSLQDPIIARMSIGSLRKSPSSCSLIASSPEENWPACQKAYN 1619

Qy       1621 LNRTPSTVTLNNNTAPTNRANQNPPEIEGIPETSQVILRPGPSPNPTAVQNENLKSMAHK 1680
              |:::|||||:.:|| : .|:| :.||| :|| :|| ||| ||
Db       1620 LNRTPSTVTLNNNSAPANRANQNPPEMEGIRETSQVILRPGSSPNPTTIQNENLKSMTHK 1679

Qy       1681 RSQRSSYTRLSKDASELH AASSESTGFGEERESD 1715
              ||||:'||||:. ||| ' |||: |||:  ||
Db       1680 RSQRSSYTRLSKDPPELHAAASSESTGFGEERESD 1715
```

RESULT 5
AAM39025
ID AAM39025 standard; Protein; 1715 AA.
XX
AC AAM39025;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2170.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy Drager syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34284.
XX
PR 21-JAN-2000; 2000US-0488[illegible].
PR 25-APR-2000; 2000US-0[illegible].
PR 09-JUN-2000; 2000US-0[illegible].
PR 19-JUL-2000; 2000US-[illegible].
PR 07-AUG-2000; 2000US-[illegible].
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang C,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue A, Yang Y, Zhang J
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT,
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI58181.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries
XX
PS Example 4; SEQ ID NO 2170; 1007pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42218) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1715 AA;

```
  Query Match             93.4%;  Score 8266;  DB 22;  Length 1715;
  Best Local Similarity   92.7%;  Pred. No. 0;
  Matches 1591;  Conservative   53;  Mismatches   70;  Indels    2;  Gaps    2;

Qy          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60
              |:|:|  :|| :|||: ||||| |||:|:|||||:|||||:|||||:||
Db          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Qy         61 GANCNLEDKDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
              |||.:|||:|||:|:|||: ||||.:| |:|:|||||| |||| :|||||.|
Db         61 GANCNLEDKDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
              .|||:. :|| '||||: ||||' |||| :|||||: |||| '||||.|'|||||
Db        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

Qy        181 KGHLECVFHLLAMGADVEQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNITDKDGNTA 240
              :|||| '|| |||| |||;' ||| '|||| |||| '||||.|'|||||
Db        181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNITDKDGNTA 240

Qy        241 LMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
              .|||:'' |.|'||'|||||:|||||||||||||||||||||:'|||||:'||||
Db        241 LMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
              |: ||| '|| :.||| :.||| : |||:|||||:|||| '.|||| .'|||
Db        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy        361 KVSAVDKKGQTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
              | |||||: |'|:|||'|||:'.|||:| |||:| |||| |'||||||||'|||
Db        361 KVSAVDKKGQTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy        421 TQIFGAPHSMTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480
              | '||| :|' ||:. ||| |||| ..|:|: .|||  | ||||' |||
Db        421 TQIFGAPHSMTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480

Qy        481 EDEMKTFAQQQIEPLFQFSWLIVFLTLLLCGGLGLVFAFPVDTNLAIAISLSFLALIYIF 540
              ||: |||  | |: |||:| |||:| |||:| '|:|| . || :|:||||||': ||
Db        481 EDEMKTFAQQQIEPLFQFSWLIVFLTLLLCGGLGLLFAFTVHPNLGIAVSLSFLALIYIF 540

Qy        541 PIVIYPGGPREGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 600
              |||||||| ||||:|||| |||:| |||||| |||:|||||||||||||||||:|||
Db        541 PIVIYPGGPREGESWNWAWVLSTRLARHIGYLELLLKLMFVNPPELPEQTTKALPVRFLF 600

Qy        601 TDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 660
              ||:|||':.||||'|||||||||||||||||||||||||||:|||:|||||||||||||
Db        601 TDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660

Qy        661 FVIFLFIVGCIIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720
              |.:|| :||| :|||'|||||, ||||'|||:'|||||||||||||||||||||.|||
Db        661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKRLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780
              |||||||:||| |||'|||||' |||||||||||||||||||:|||||:|||| |||
Db        721 NSQRKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy        781 EQDKVLQMIDTVRVLFSKGPFIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              ||| ||  |:|:|||||:|||||:||| |||||||||||||||||||:|||| |||
Db        781 EQDKVLQMIDTVRVLFSKGPFIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
```

```
Qy        841 LPVFLNSRGLSNAEKFLVTSAINGGITCSDTTIQEDTDRRVSQNSLIEMTKLISKTAIN 900
              ||||||||: |||||  : ||||   :  |:  |    | .  ||.  ||  ||  ||
Db        841 LPVFLNSRGLSNAEKFLVTSATNGDVPCSDITIQEDADRRVSQNSLIEMTKLISKTAIN 900

Qy        901 RRDTYRRRQMQRTITRQMSFDLTKLLVTEEWF  ||||  |IMPPLLNIVSVTGRILRANQIT 960
              ||  ||||  ||||  ||||  ||  |  |  |  ||  ||  ||  ||  ||  ||
Db        901 RRDTYRRRQMQRTITRQMSFDLTK LVTEDWF  ||||  |IMPPLLNIVSVTGRILRANQIS 960

Qy        961 FNWDRLASWINLTEQWPYRTSWLIQYLEETEGVEFMYFERIPPNIPTTKDVREIL 1015
              |||  |  |||  |  ||||  |  ||  |  ||  |  ||  ||  |  ||  ||
Db        961 FNWDRLASWINLTEQWPYRISWLIQYLEETEGVEFMYFERIPPNIPTTKDVREIL 1015

Qy       1016 EIDGDIRNPEVFLPSPTPVLVAPIVKVPLPGTTNSIDPEIPEIIADVRAARPEQINIGSLAY 1075
              |||  ||||  |  ||  |||  |  |  ||  ||  |  ||  |  ||  ||  ||
Db       1016 EIDGDIRNPEVFLPSPTPVLVAPIVKVPLPGTTNSIDPEIPEIIADVRAARPEQISIGSLAY 1075

Qy       1076 PPLPLHEGPPRPFSGYSQFASVTSNASFNGFPTTGVSVPQFHSSYYSSLSGPQHFFYNRA 1135
              ||  ||||  |  ||  |  ||  |  ||  |  ||  ||  ||  ||  ||  ||  ||
Db       1076 PPLPLHEGPPRPASGYSQFPSVTSNTSFNGFPATGVSVPQFHSSYYSSMTGPQHFFYNRS 1135

Qy       1136 AVPATGSSLNNSSMTVDVVTEKDPIEGLDQNPMFQVTTIPKANINGRVLSQNIDELR 1195
              :  |    :|  :  ||    |:  |  ||  :  |  ||  ||  :|  :|  ||
Db       1136 SGPAPGPVVLLNSTNVDAVTEKDPIEGLDQSPMIPQVTTIPKANINGRVLAQNIDELR 1195

Qy       1196 KEMAMNPQDWHLFPSMVLEMRPVEMQVVPEDTPFLNPAVSSAPVPHSEPARRSNHTELPTT 1255
              |||  ||||  |  ||  |  ||  |  ||  |  ||  ||  ||  ||  ||  |  |
Db       1196 KEMNNNPGDWHLFPSTVLEMRNAEMHVVPEDTPFLNPSVSSGPVAIHGPPAFRASHNELPHT 1255

Qy       1256 ELSSQTPYTLNPSPEELNTIGIDPIAPPHSNLWQAQIERPTPSLSSINSQDSNIEISKIT 1315
              .||  |  ||||  |  ||  |  ||  ||  ||  ||  ||  ||  ||  ||  ||  |
Db       1256 ELSSQTPYTLNPSPEELNTIGIDPIAPPHSNIWQSQTERPTPSLSSINSQDSNIEISKIT 1315

Qy       1316 DKVQAEYRDAYREYIAQMSQLEQRYGSTTISSSRFHQTYYTQQSSSRFSIHHGLEQEHQ 1375
              |||  ||||  |||  ||  |  |  |  |  |  |  |  |  |  ||  ||  ||
Db       1316 DKVQAEYRDAYREYIAQMSQLEQRYGSTTISSSRFHQTYYMQQSSSRGSIHSNLEQEHQ 1375

Qy       1376 KEGKILKQEDGRKSFLMKRGDVIDYSSSGVSTNRASPLDPITEEDEPSDQSGSKLDPGKFS 1435
              |  |  :|||  ||||  ||  |  ||  ||  ||  ||  ||  ||  ||  ||  |
Db       1376 KDSEPKPDDGRKSFLMKRGDVIDYSSSGVSINRASPLDPITEEDEKSDQSGSKLDPGKFS 1435

Qy       1436 SERPSLFQTDLKLPGGGLAYQKLIDEDESMT RVQITTHTNKMIETKPLKAAQRECANP 1495
              ||  |  :|||  ||  |  ||  |  |||  |  |  |  |  |  |  |  |  |
Db       1436 SERSSLFQTDLKLPGGGLAYQKLIDEDESMT KVQITTHTNKMIETKPLKAAQRECANP 1495

Qy       1496 QEHLAEPIPIPIPYPPYPLLAPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIP 1555
              :||||  ||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1496 PEHVAEPIPIPIPAPPYPLLAPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIP 1555

Qy       1556 ELETEGHSGKPQMEDSSSQLPPARMLPPPPPPPPPPPPPPPPPPPPESWSSPLKAVN 1615
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1556 ELETDGHSGKRGIPPSLSGLQDPTIARMLPPPPPPPPPPPPPPPPASSPEENWIACPKAVN 1615

Qy       1616 LNRTPSTVTLNNNTAPTNRANQNPDLEIEDIRFLQVLYEGSCNFTAVQNENLKSMAHK 1675
              ||  |||  |  ||  |  |||  |  ||  |  ||  ||  |  ||  ||  ||
Db       1616 LNRTPSTVTLNNNSAPANRANQNPDMEMEGIRSLQVLYEGSCNFTTIQNENLKSMTHK 1675

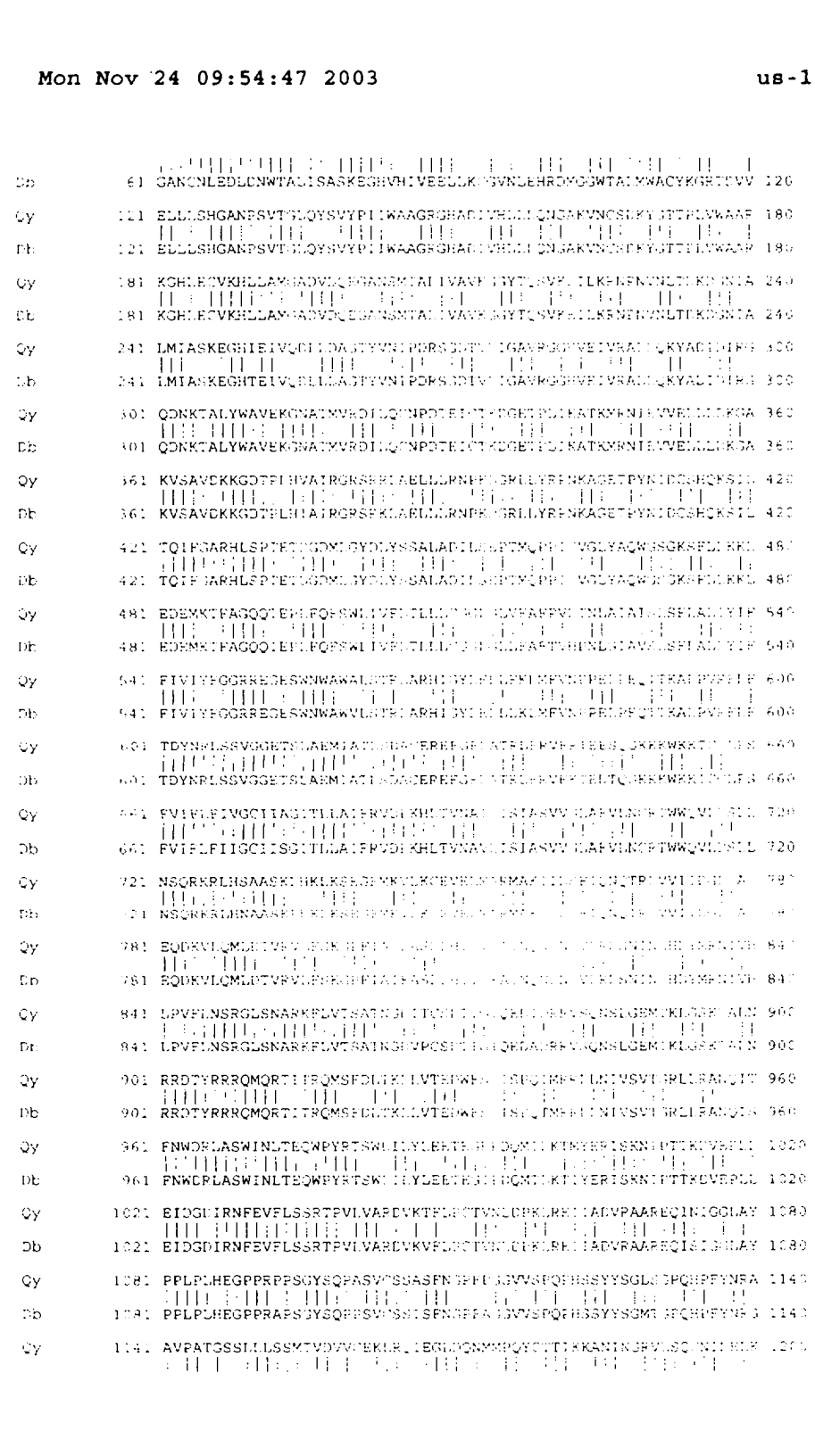
Qy       1676 RSQRSSYTRLSKDASELH-AASSEATGPHEEPPPPPLL 1712
              |||||||:|||  |  |  ||||  |||  |
Db       1676 RSQRSSYTRLSKDPFELHAAASSRATGPHEEPPPPPLL 1712
```

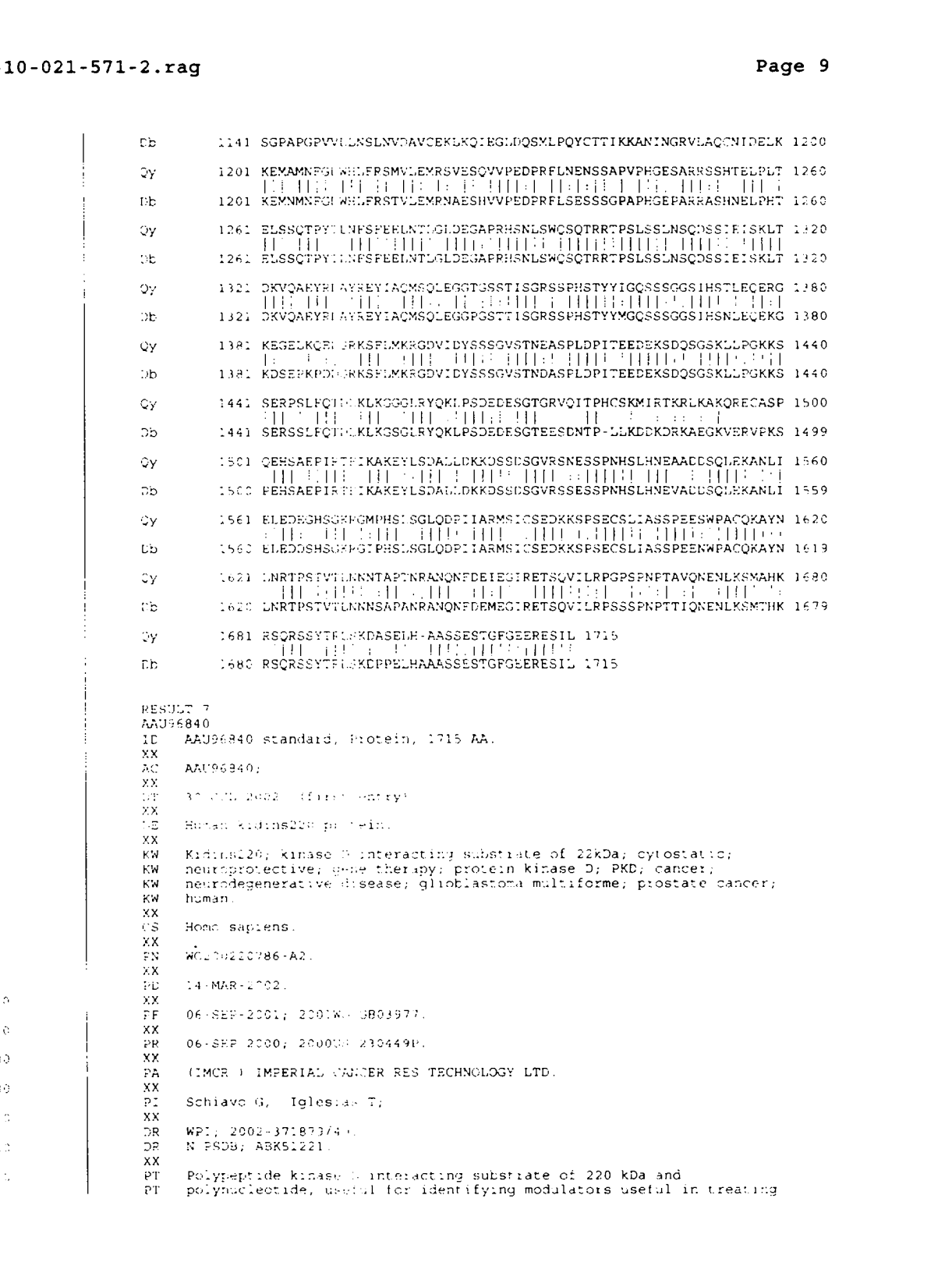
RESULT 6
AAE25144
ID AAE25144 standard; Protein; 1715 AA.
XX
AC AAE25144;
XX
DT 30-OCT-2002 (first entry)
XX
DE Human ARMS protein.
XX
KW Ankyrin repeat rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW diagnostic; imaging; human.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Region   1..390
FT   /note= "Ankyrin repeat"
FT Region   ...493
FT   /note= "N terminal region"
FT Domain   ...518
FT   /note= "Transmembrane domain"
FT Domain   ...546
FT   /note= "Transmembrane domain"
FT Region   ...639
FT   /note= "Loop between TM2 and TM3"
FT Domain   ...680
FT   /note= "Transmembrane domain"
FT Domain   ...710
FT   /note= "Transmembrane domain"
FT Region   749..854
FT   /note= "Carboxy tail"
FT Region   940..1060
FT   /note= "Carboxy tail"
FT Region   1081..1094
FT   /note= "Polyproline stretch"
FT Domain   1151..1221
FT   /note= "SAM domain"
FT Binding site   1713..1715
FT   /note= "PDZ binding motif"
XX
PN WO200250273-A2.
XX
PD 27-JUN-2002.
XX
PF 19-DEC-2001; 2001WO-US48603.
XX
PR 21-DEC-2000; 2000US-256909P.
XX
PA (UYNY) UNIV NEW YORK STATE.
XX
PI Chao MV, Kong H;
XX
DR WPI; 2002-608800/64
DR N-PSDB; AAD41036
XX
PT Isolated ankyrin repeat rich membrane spanning (ARMS) polypeptide that
PT is a target for phosphorylation by neurotrophin and ephrin receptor
PT tyrosine kinases, useful as a marker for growth cones -
XX
PS Claim 1; Page 109-110; 136pp; English.
XX
CC The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
CC protein which is a down stream target of neurotrophin and ephrin
CC receptor tyrosine kinases. ARMS DNA is useful for visualising the
CC growth cone of neurons. ARMS protein is useful as an indicator of
CC the biological activity of neurotrophins and ephrins, as a marker
CC for neuronal cells which have the ability to undergo continued
CC synaptic changes through adult life or for the presence and
CC distribution of ARMS in such neuronal cells, or as a marker for
CC growth cones. ARMS DNA is useful in diagnostic and imaging methods.
CC The present sequence is human ARMS protein.
XX
SQ Sequence 1715 AA;

  Query Match   93.4%; Score 8266; DB 23; Length 1715;
  Best Local Similarity   92.7%; Pred. No. 0;
  Matches 1591; Conservative   53; Mismatches   70; Indels   2; Gaps   2;

```
Qy          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60
              ||:|||  |||||||||||||||:|||||:|||||||||||||||:||||:|||:||||
Db          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Qy         61 GANCNLEDKDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
```





RESULT 7
AAU96840
ID AAU96840 standard; protein; 1715 AA.
XX
AC AAU96840;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human kidins220 protein.
XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer;
KW human.
XX
OS Homo sapiens.
XX
PN WO200210786-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000US-230449P.
XX
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Schiavo G, Iglesias T;
XX
DR WPI; 2002-371873/41.
DR N-PSDB; ABK51221.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating

PT cancer and neurodegenerative diseases
XX
PS Claim 33; Fig 11; 228pp; English.
XX
CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by GenBank Accession No. AB033766. Also included are
CC a polynucleotide encoding Kidins220 provided that polynucleotide is
CC not any one of the clones corresponding to that of GenBank Accession Nos.
CC given in the specification; an expression replicable vector comprising
CC the polynucleotide; a recombinant host cell containing the polynucleotide
CC or vector; an anti-Kidins220 antibody useful in the preparation of
CC Kidins220; an agent (A1) which modulates activity of Kidins220,
CC protein kinase D (PKD) or interaction between PKD and Kidins220;
CC a mutant animal transgenic for kidins220; the use of an agent capable of
CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
CC reagent for diagnosing or prognosing cancer or for monitoring the
CC progression of cancer in a patient; a compound comprising a moiety that
CC selectively binds to Kidins220 or its variant and another moiety, and a
CC nucleic acid molecule encoding the compound. The antibody is useful for
CC modulating the activity of Kidins220, where the antibody prevents a
CC region of Kidins220 interacting from another protein, or reduces the
CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
CC Kidins220 polynucleotide is useful for identifying an agent which
CC modulates the promoter activity of the polynucleotide, and Kidins220 is
CC useful for identifying an agent which modulates activity of Kidins220,
CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
CC Kidins220 or its polynucleotide is useful to evaluate, e.g. using gene
CC therapy or for diagnosing or modulating progression of cancer in a
CC patient. The agent is useful in manufacture of a medicament for use in
CC treatment of neurodegenerative disease, and the agent which inhibits
CC function of Kidins220 gene or its product is useful in treating cancer
CC which is glioblastoma multiforme, a prostate cancer. The compound is
CC useful for imaging cancer in an individual, and in diagnosing or
CC prognosing, and also for treating an individual. The present
CC sequence represents human kidins220.
XX
SQ Sequence 1715 AA;

Query Match 94.4%; Score 8...; ... ...; Length 1715;
Best Local Similarity 91.7%; Pred. No. 0.;
Matches 1590; Conservation ...%; Mismatches ...; Indels ...; Gaps ...;

```
Db       1500 PEHSAEPIRTPFIKAKEYLSDALLDKKDSSDSGVPSSESSPNHSLHNEVAIDSQLEKANLI 1559

Qy       1561 ELEDEGHSGKRGMPHSLSGLQDPIIARMSIDSLIKRSPSEGSLIASSPEESWPACQKAYN 1620
              |.|.| || | |.|| .|||. .|||| |||  ||  ||  |||. |||  | |

Db       1560 ELEDDSHSGKPGIPHSLSGLQDPIIARMSIDSLIKRSPSEGSLIASSPEESWPACQKAYN 1619

Qy       1621 LNRTPSTVTLNNNTAPTNPANQNFIEIESIPRTQVIIRPGSPSTATVLNENIKSKAHF 1680
              |||||. |||| .|| .|| . .|| .|| |  |  |  |  |  |  |  |  |

Db       1620 LNRTPSTVTLNNNSAPANRANQNFIEMESIPRTQVIIRPGSSSIATTLNENIKSKIHK 1679

Qy       1681 RSQRSSYTRLSKIASRLH AAASFSTGPGEERESII 1715
              ||||.| ||||    |  |||  ||||  |||

Db       1680 RSQRSSYTRLSKIPRLHAAAASFSTGPGEERESIL 1714


RESULT 6
AAB32128
ID   AAB32128 standard; Protein; 1715 AA.
XX
AC   AAB32128;
XX
DT   24-MAR-2003 (first entry)
XX
DE   Human cytoskeleton-associated protein, CSAP 26.
XX
KW   Human; cytoskeleton associated protein, CSAP 26; atherosclerosis,
KW   cancer; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          42..96
FT                   /note= "Domain present in CSAP 1"
FT   Domain          354..392
FT                   /note= "Domain present in CSAP"
FT   Binding site    467..474
FT                   /note= "ATP/GTP-binding site motif A"
FT   Domain          494..514
FT                   /note= "Transmembrane domain"
FT   Domain          524..544
FT                   /note= "Transmembrane domain"
FT   Domain          654..674
FT                   /note= "Transmembrane domain"
FT   Domain          687..707
FT                   /note= "Transmembrane domain"
XX
PN   WO200170404 A.
XX
PD   10-OCT-2001.
XX
PF   25-MAR-2002; 2002WO-US009269.
XX
PR   29-MAR-2001; 2001US-280606P.
PR   03-APR-2001; 2001US-281323P.
PR   13-APR-2001; 2001US-283769P.
PR   24-MAY-2001; 2001US-289609P.
PR   10-MAY-2001; 2001US-290518P.
PR   18-MAY-2001; 2001US-291870P.
PR   29-MAY-2001; 2001US-294451P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hafalia AJA, Tang TY, Yue H, Khan FA, Isom CH, Baughn MR;
PI   Warren BA, Duggan BM, Thangavelu K, Honchell CD, Azimzai Y;
PI   Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
PI   Richardson TW, Lee SY, Bandman J, Lal PG, Lee S, Gietzen KJ;
PI   Walia NK, Griffin JA, Lee EA, Swarnakar A, Rink HE, Jones KA;
XX
DR   WPI; 2003-092694/08.
DR   N-PSDB; AAD49615.
XX
PT   New human cytoskeleton associated proteins useful for preparing a
PT   composition for diagnosing or treating a disease or condition

PT   associated with decreased expression or overexpression of functional
PT   CSAP e.g., cancer.
XX
PS   Claim 1; Page 195 149; 233pp; English.
XX
CC   The invention relates to new human cytoskeleton-associated protein
CC   (CSAP) and its polynucleotide. The polypeptide is useful for preparing
CC   a composition for diagnosing or treating a disease or condition
CC   associated with decreased expression or overexpression of functional
CC   CSAP e.g. atherosclerosis or cancer. The present sequence is human
CC   CSAP-26 protein. The invention is useful in gene therapy.
XX
SQ   Sequence   1715 AA,

  Query Match             93.4%;  Score 8266;  DB 24;  Length 1715;
  Best Local Similarity   92.7%;  Pred. No. 0;
  Matches 1591;  Conservative   53;  Mismatches   70;  Indels    2;  Gaps    2;

Qy          1 MSVLISQSVINYVEEENIPALKALIEKCKDVDERNECGQTPLMLAAEQGNVEIVKELIKN 60
              || |||   |||| |||  |||| |||| | ||||.||||..||||.:.|||||||  |
Db          1 MSVLISQSVINYVEEENIPALKALIEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Qy         61 GANCNLEDKFNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
              |||. |||   ||| |||. |:|:' |||| : :|||.|||||' ||||||:: ||||
Db         61 GANCNLEDKFNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
              |||||:|.  ||| ||||  |||| .|||| : |||||||||.' |||||  ||
Db        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

Qy        181 KGHLECVKLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
              |||.:||.  ||| ||||  |||| ||||||||||||||||:|||||  ||||| ||
Db        181 KGHLECVKLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy        241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
              ' ||| ' :.:'|||.:|||' |||| :|||||:|||||||'||||||:|||||||||
Db        241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
              ||||:.|||| ||||  |||||||||||||||||||||||'||||||||||||||||||
Db        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy        361 KVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
              |:|||   |:' :|:.|||.:|||| |||| |||||:|||||||||||||||||||||
Db        361 KVSAVDKKGETPLHIAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy        421 TQIFGASPFSIETDGIMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480
              |.' |||   | :'|||  ||| '|'| : |||||:|||| '.||||| . ||||
Db        421 TQIFGASPFSIETDGIMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480

Qy        481 EDEMKTFAVSLIEPLFQFSWLIVFLTLLLCGGLGLVFAFPVDTNLAIAISLSFLALIYIF 540
              || :|||, : | .'||| |||| ||||:: |:||. ' || |:|||||||:||
Db        481 EDEMKTFAVSLIEPLFQFSWLIVFLTLLLCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540

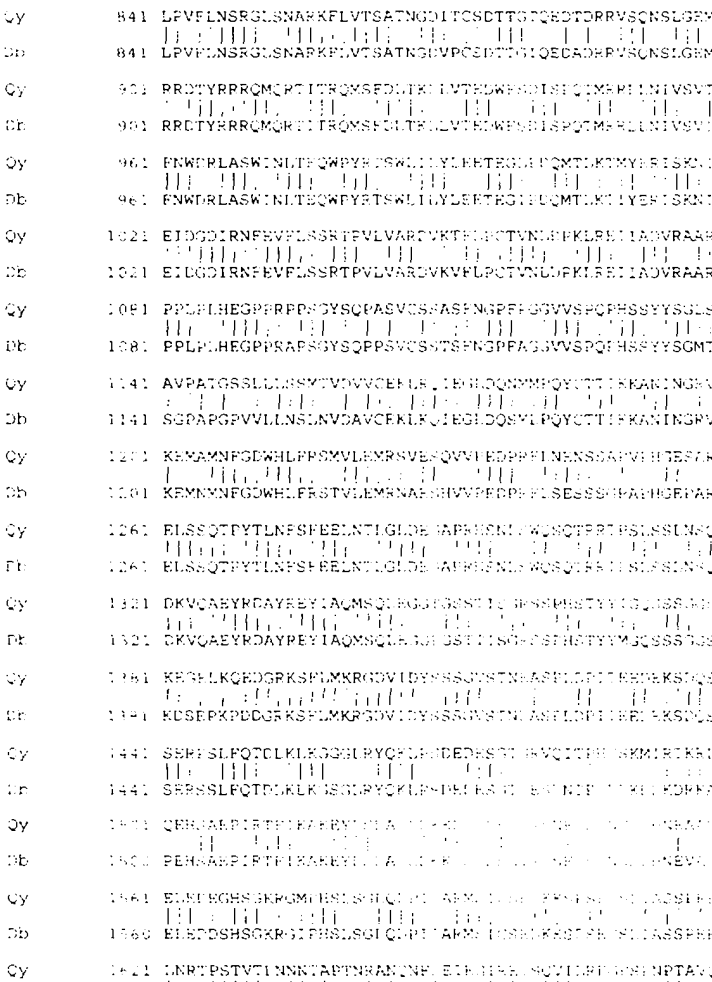
Qy        541 FIVIYFGGPREGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 600
              |||'|||  |||'.|||  |||| |||||:|| |:|||||'|||||:|.||||
Db        541 FIVIYFGGPREGESWNWAWVLSTRLARHIGYLELLLKLMFVNPPELPEQTTKALPVRFLF 600

Qy        601 TDYNRLSSVPGETSLAEMIATLSDACEREFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 660
              |||. ||| '|||||||||:|||':||||||||||||||||:|::|||||||||||||
Db        601 TDYNRLSSVPGETSLAEMIATLSDACEREFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660

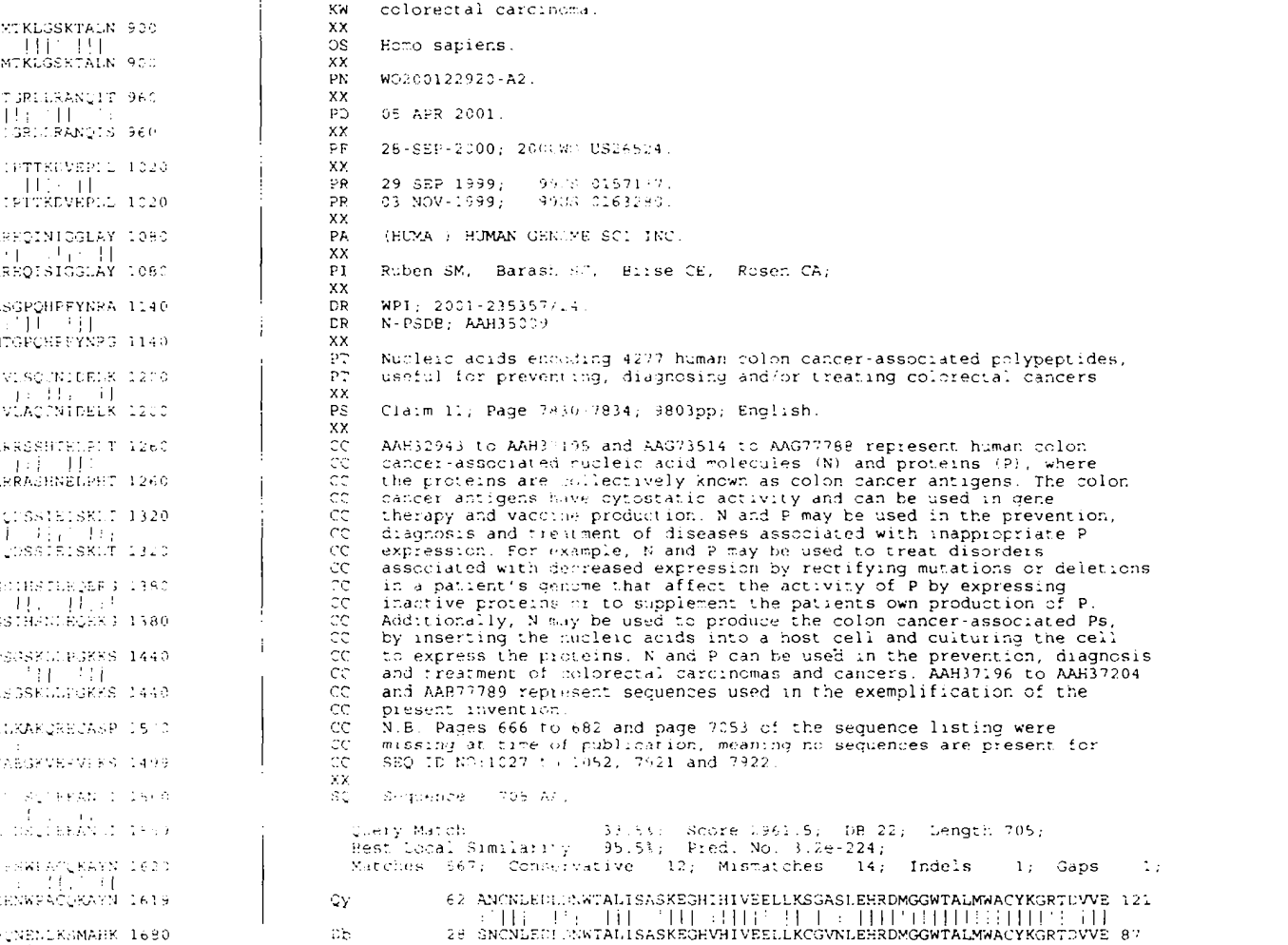
Qy        661 FVIFLFIVGIIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720
              |||||||:'||:||||.|'|||'||||||'|:|||||||||||||'||||||||
Db        661 FVIFLFIIGIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKRLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780
              |||||||.' |||'|||.|'|||:'||||||||||||||||:'||||||||||||||
Db        721 NSQRKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy        781 EQDKVLQMIIVRVLFSKGPFIAIFASDFHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              |||' ||| |||:' |||'|||| |||||:|||||:|||:|||||'||||
Db        781 EQDKVLQMIIVRVLFSKGPFIAIFASDFHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
```

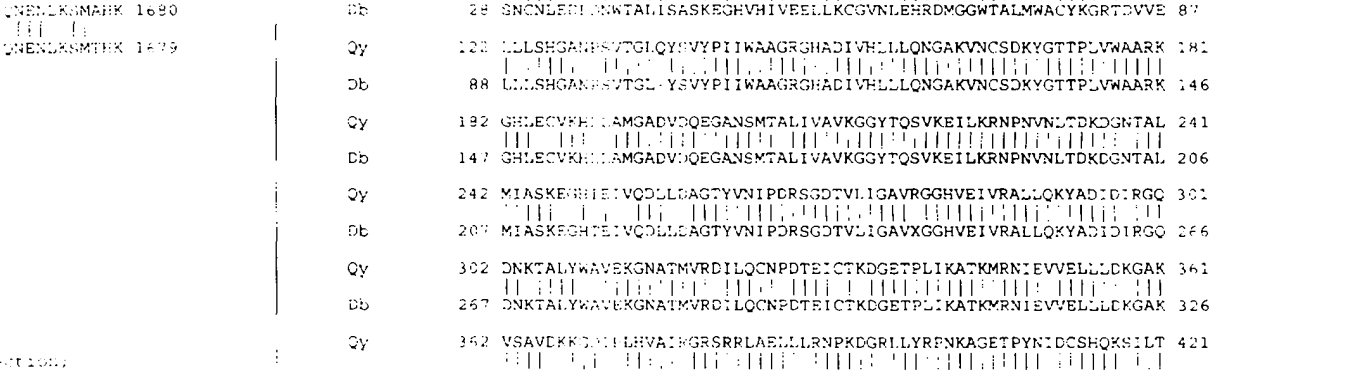


RESULT 9
AAG75604
ID AAG75604 standard; Protein; 705 AA
XX
AC AAG75604;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:6368
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;



KW colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05 APR 2001.
XX
PF 26-SEP-2000; 2000WO-US26504.
XX
PR 29 SEP 1999; 99US-01571-07.
PR 03 NOV-1999; 99US-0163280.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH35009.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers
XX
PS Claim 11; Page 7430-7834; 9803pp; English.
XX
CC AAH32943 to AAH37196 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 [illegible] 1052, 7921 and 7922.
XX
SQ Sequence 705 AA;

Query Match 33.5%; Score 1961.5; DB 22; Length 705;
Best Local Similarity 95.5%; Pred. No. 3.2e-224;
Matches 567; Conservative 12; Mismatches 14; Indels 1; Gaps 1;



```
Db        327 VSAVDKKGDTXLHIAIRGRSRKLAELLLRNPKDGRLLYRENKAGETPYNIDCSHQKSILT 386

Qy        422 QIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKLE 481

Db        387 QIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKLE 446

Qy        482 DEMKTFAGQQTEPLFQFSWLIVFLTLLLCGGLGLVFAFPVFILNLAIASLSFLALIYIFF 541

Db        447 DEMKTFAGQQIEPLFQFSWLIVFLTLLLCGGLGLVFAFTVFILNLSIAVRSFLALLYIFF 506

Qy        542 IVIYFGGRRREGESWNWAWAISTSLARHIGYLEIFLKLVFVNSFFIPSQITFALPVFFLFT 601

Db        507 IVIYFGGRREGESWNWAWVISTSLARHIGYLEIFLKLVFVNSFFIPSQITFALPVFFLFT 566

Qy        602 DYNRLSSVGGETSLAEMIATLSDAIERFEGFLATRLFRVFFIFRSQDKKKWKKT 655

Db        567 DYNRLSSVGGETSLAEMIATLSDAIERFEGFLATRLFRVFFIFRSQDKKKKKNS 620
```

```
RESULT 10
AAE01035
ID   AAE01035 standard; Protein; 551 AA.
XX
AC   AAE01035;
XX
DT   04-APR-2001 (first entry)
XX
DE   Human death domain-containing receptor (DDCR) protein from HODDX59 clone.
XX
KW   Human; death domain-containing receptor; DDCR; immunosuppressive;
KW   antiarthritic; antirheumatic; antiproliferative; cytostatic;
KW   cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW   antibacterial; virucide; fungicide; ophthalmological; gene therapy;
KW   immunodeficiency disease; Acquired immune deficiency syndrome;
KW   AIDS; leukaemia; autoimmune disease; systemic lupus erythematosus;
KW   hyperproliferative disorder; neoplasm; cerebrovascular disorder;
KW   cerebral ischaemia; angiogenesis; cardiovascular disorder;
KW   neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW   ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;
KW   spinal muscular atrophy; epithelial cell proliferation; infection;
KW   cancer; wound healing; skin aging; chemotaxis; HODDX59 clone.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          599..615
FT                   /label= Transmembrane domain
FT   Domain          626..644
FT                   /label= Transmembrane domain
FT   Region          64..69
FT                   /label= Immunogenic_epitope
FT   Region          131..136
FT                   /label= Immunogenic_epitope
FT   Region          164..169
FT                   /label= Immunogenic_epitope
FT   Region          197..203
FT                   /label= Immunogenic_epitope
FT   Region          229..235
FT                   /label= Immunogenic_epitope
FT   Region          298..316
FT                   /label= Immunogenic_epitope
FT   Region          328..336
FT                   /label= Immunogenic_epitope
FT   Region          445..452
FT                   /label= Immunogenic_epitope
FT   Region          482..490
FT                   /label= Immunogenic_epitope
FT   Region          540..551
FT                   /label= Immunogenic_epitope
XX
PN   WO200129063-A2.
XX
PD   26-APR-2001.
XX
PF   17-OCT-2000; 2000WO-US28666.
XX
PR   18-OCT-1999;   99US-0159565.
PR   24-NOV-1999;   99US-0167246.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J, Ruben SM;
XX
DR   WPI; 2001-300316/31.
DR   N-PSDB; AAD04776.
XX
PT   New death-domain containing receptor polynucleotides and polypeptides,
PT   useful for treating and diagnosing cancer -
XX
PS   Claim 1; Page 292-294; 298pp; English.
XX
CC   The patent discloses novel death domain-containing receptor (DDCR)
CC   cDNAs and their corresponding proteins. DDCR cDNA and protein are
CC   used to prevent, treat or ameliorate a medical condition in mammals.
CC   They are also used in diagnosing a pathological condition or
CC   susceptibility to a pathological condition. The DDCR protein and
CC   its antibodies are used in the diagnosis and treatment of disorders
CC   such as immunodeficiency diseases (e.g. Acquired immune deficiency
CC   syndrome (AIDS), leukaemia) autoimmune diseases (e.g. systemic lupus
CC   erythematosus, rheumatoid arthritis), hyperproliferative disorders
CC   (e.g. neoplasms of the breast or liver), cerebrovascular disorders
CC   (e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders
CC   (e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's
CC   disease, Parkinson's disease), ocular disorders (e.g. corneal
CC   infection), degenerative diseases (e.g. spinal muscular atrophy-SMA),
CC   cancer, aberrant apoptosis, disorders of the placenta or uterus and
CC   infections caused by bacteria, viruses and fungi. The DDCR proteins
CC   are used to aid wound healing and epithelial cell proliferation,
CC   to prevent skin aging due to sunburn, to maintain organs before
CC   transplantation, to support cell culture of primary tissues, to
CC   regenerate tissues and in chemotaxis. The DDCR sequences are used
CC   in correcting aberrant cellular apoptosis by gene therapy.
CC   The present sequence is human death domain-containing receptor
CC   (DDCR) protein from clone HODDX59.
XX
SQ   Sequence   551 AA,

  Query Match             31.4%;  Score 2782.5;  DB 22;  Length 551;
  Best Local Similarity   96.7%;  Pred. No. 2.8e-210;
  Matches  534;  Conservative    8;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        102 MGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNG 161

Db          1 MGGWTALMWACYKGRTDVVELLLSHGANPSVTGL YSVYPIIWAAGRGHADIVHLLLQNG 59

Qy        162 AKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVK 221

Db         60 AKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVK 119

Qy        222 EILKRNPNVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 281

Db        120 EILKRNPNVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 179

Qy        282 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 341

Db        180 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 239

Qy        342 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRP 401

Db        240 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRP 299

Qy        402 NKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 461

Db        300 NKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 359

Qy        462 CVGLYAQWGSGKSFLLKKLEDEMKTFAGQQTEPLFQFSWLIVFLTLLLCGGLGLVFAFPV 521
```

```
Db        360 CVGLYAQWGSGKSPLLKKLEDEMKTPA QQIEI FQPSWLIVFLTL  QHSDGLLFAFIV 419

Qy        522 DTNLAIAISLSFLALLYIFFIVIYF GRRREGE WNWAWAL TRLASHI  LEI LFKLMFV 581

Db        420 HPNLGIAVSLSFLALLYIFFIVIYF GRREGE WNWAWV  TRLASHI YLEL LFKLMFV 479

Qy        582 NPPELPEQTTKALPVRFLFIDYNP          AEMIAT  DA  RRE  LATRIFFUF 641

Db        460 NPPELPEQTTKALPVRFLFIDYNP  SQV  ET LAEMIAT  DA  RRE  LATRIFFUF 539

Qy        641 RTEESQGKKKWK 653

Db        540 KTEDTQGKKKKK 551
```

RESULT 11
AAB95191
ID   AAB95191 standard; Protein; 545 AA.
XX
AC   AAB95191;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human protein sequence SEQ ID NO:17273.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, [illegible]
PI   Ishii S, Sugiyama T, Wakamatsu A, [illegible]
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full length cDNAs -
XX
PS   Claim 8; SEQ ID 17273; 2537pp + [illegible] English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides, or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5' end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH19742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   545 AA;

```
  Query Match             27.2%;  Score 2408;  DB 22;  Length 545;
  Best Local Similarity   86.8%;  Pred. No. 1e-180;
  Matches  472;  Conservative   24;  Mismatches   46;  Indels    2;  Gaps    2;

Qy       1173 MMPQYTTIKKANINGRVLSQCNIDELKKEMAMNFGDWHLFRSMVLEMRSVESQVVPEDP 1232

Db          1 MLPQYTTIKKANINGRVLAQCNIDELKKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDP 60

Qy       1233 RFLNENSSAPVPHGESARRSSHTELPLTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 1292

Db         61 RFLSESSSPAPHGEPARRASHNELPHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 120

Qy       1293 SWQSQTRRTPALSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGTGSSTISG 1352

Db        121 SWQSQTRPIPSLSSINSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISG 180

Qy       1353 RSSPHSTYYMSQSSSGGSIHSTLEQERGKEGELKQEDGRKSFLMKRGDVIDYSSSGVSTN 1412

Db        181 RSSPHSTYYMSQSSSGGSIHSNLEQEKGKDSEPKPDDGRKSFLMKRGDVIDYSSSGVSTN 240

Qy       1413 EASPLDPITEEDEKSDQSGSKLLPGKKSSERPSLFQTDLKLKGGGLRYQKLPSDEDESGT 1472

Db        241 DASPLDPITEEDEKSDQSGSKLLPGKKSSERSSLFQTDLKLKGSGLRYQKLPSDEDESGT 300

Qy       1473 GRVQIIDPNEMIRTKRLKAKQRECASPQEHSAEPIRTFIKAKEYLSDALLDKKDSSDSG 1532

Db        301 EESDNTP LLKDDKDRKAEGKVERVPKSPEHSAEPIRTFIKAKEYLSDALLDKKDSSDSG 359

Qy       1533 VRSNESSPNHSLHNEAADDSQLEKANLIELEDEGHSGKRGMPHSLSGLQDPIIARMSICS 1592

Db        360 VRSSESSPNHSLHNEVADDSHLEKADLIELEDDSHSGKRGIPHSLSGLQDPIIARMSICS 419

Qy       1593 EDKKSPSE SLIASSPEESWPACQKAYNLNRTPSTVTLNNNTAPTNRANQNFDEIEGIRE 1652

Db        420 EDKKSPSE SLIASSPEENWPACQKAYNLNRTPSTVTLNNNSAPANRANQNFDEMEGIRE 479

Qy       1653 TSQVIISE PNPTAVQNENLKSMAHKRSQRSSYTRLSKDASELH-AAASESTGFGEER 1711

Db        480 TSQVIISE SNPTTIQNENLKSMTHKRSQRSSYTRLSKDPPELHAAASSESTGFGEER 539

Qy       1712 ESIL 1715

Db        540 ESIL 543
```

RESULT 12
ABB64857
ID   ABB64857 standard; Protein; 1498 AA.
XX
AC   ABB64857;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 21363.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.

```
XX
PR   03-MAR-2000; 2000US-191637P.
PR   11-JUL 2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC, Adams M, [illegible]
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL08960.
XX
PT   New isolated nucleic acid detection reagent for detecting 1 [illegible] or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions
XX
PS   Disclosure; SEQ ID NO 21363; 2[illegible]pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL14176-ABL3[illegible]), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1498 AA;

  Query Match             26.8%;  Score [illegible];  DB 22;  Length 1498;
  Best Local Similarity   37.6%;  Pred. No. [illegible];
  Matches  552;  Conservative  2[illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps   36;

Qy          1 MSVLISQSVINYVEHENIFALFA[illegible]KHELLK 59
Db         23 MGSLGHRALLQYI[illegible]NDISGLPAI[illegible]FLA 82

Qy         60 NGANCNLEDLDNWTALISACKE[illegible]PTPV 119
Db         83 RGADVQAEDLDNWTALLCASEN[illegible]HIRL 142

Qy        120 VELLLSHGANP[illegible]VWAA 179
Db        143 VRLLL[illegible]

Qy        180 RK[illegible]
Db        202 RRGNVEIVDTLL[illegible] 261

Qy        240 ALMIASKEGHIEIVQDLLDAG[illegible]KYAD[illegible] 299
Db        262 ALCIASREGHQDIAASLIAAGAYIN[illegible]ALLKKHADVDIQ 321

Qy        300 GQDNKTALYWAVEKGNATMVRD[illegible]KATKMPNIEVVELLL[illegible] 359
Db        322 GKDRKTAIYTAVEKGHTPIVKL[illegible]VHLLLDRK 381

Qy        360 AKVSAVDKKGDTP[illegible]HVAIRGRSR[illegible]NIDCSHQKSI 419
Db        382 AKVTASDKRGDTCLHIAMPARSK[illegible]ETPYNIDSLHQKTI 441

Qy        420 LTQIFGARHLSPTETDGDM[illegible]YAQWGSGKSFLLKK 479
Db        442 LGQVFGARRLNTNEDSEGM[illegible]YAKWGSGKSFLLNK 501

Qy        480 LEDEMKTFAGQQTEPLFQFSW[illegible]SLSFLAL[illegible]Y 538
Db        502 LRDEMNNFARQWAEPPIRTSGL[illegible]SAAV[illegible]FLLLAY 561

Qy        539 IFFIVIYFGGRREGESWNWAWALSTR[illegible]ELPEQTTKALPVRF 598
Db        562 LLLAAVRY[illegible]NYQ [illegible]KQWAYSVQHGLEKRMTRLRLILQVAFCHPPG-PQSDSQAKPVRF 618

Qy        599 LPTDYNR[illegible]SSVGGETSLAEMIATLSDACEREFGFLATRLFRVFRTE--ESQGKKKWKKTC 656
Db        619 HPAEANSA[illegible]GDGAVAHMLAALLDAIESHYGWLATRLYRAFRPKCLKVDVGWRWRRMC 677

Qy        657 CLPSFVIH[illegible]VGCIIAGITLLAIP-----RVDPKHLTVNAILISIASVVGLAFVLNCRT 711
Db        678 CIPIV[illegible]ALVTVV[illegible]ISLTVAYFTFADEKEKEHILV--ALYVIAAVMGTLICTHLHV 735

Qy        712 WWQVL[illegible]QRKR[illegible]HSAASKLHKLKSEGFMKVLKCEVELMARMAKTIESFTQNQTRLV 771
Db        736 LAKVF[illegible]HIRVLFRAV - RSSESAPLTMLGAEVAVMTDMVKCLDAFTNQQSRLV 791

Qy        772 VIIDGLDA[illegible]QKVLQMLDTVRVLFS -KGPFIAIFASDPHIIIKAINQNLNSVLRDSNI 829
Db        792 GVIDALDS[illegible]TERILTLLNAVQTLLSSPNRPFVLLISVDPHVIAKAAEANSRRLFTEGGI 851

Qy        830 NGHDYMRN[illegible]VHLPVFLNSRGLSNARKFLVTS---  ATNGDITCSD--TTGTQEDTDRRV 882
Db        852 GGHDFL[illegible]VHLPVYLQNSGLRKVQRAQMTALLFKRSGGGDYQTDDGPTLG HSVSARRL 910

Qy        883 SQNS- LGE[illegible]TKL---  --GSKTALNR-- --RDTYRRRQMQRTITRQMSFDLTKL 925
Db        911 SNASE[illegible]QEKLFGPARGGGGKKLRLSESVASSTGSNLHRLGQNPQTV ---LDLSRI 965

Qy        926 LVTEDWFS[illegible]FQTMRRLLNIVSVTGRLLRANQITFNWDRLASWINLTEQWPYRTSWLIL 985
Db        966 VLTDDYFS[illegible]VNPRSMRRIMNVIYITVRLLKAFQIEFSWYRLSSWINLTEQWPLRASMIVL 1025

Qy        986 YLEE- TE[illegible]QMTLKTMYERISKNIPTTKDVEPLLEIDGDIRNFEVFLSSRTPVLVAR 1043
Db       1026 HHDQFMDSNADESVSLQSVYEKLRPKLAYLREAAPLLELDRDERKLDAFLQLHKSDLLVA 1085

Qy       1044 DVKTFLP[illegible]VNLDPKLREIIADVRAAREQINIGGLAYPPLPLHEGP--- PRPPSGYSQP 1099
Db       1086 DLRIFLPF[illegible]NLDPYLRKVLKEDQQTIE--DEGSLVIQARPSVSNTMRQFPAPTTYVPSP 1143

Qy       1100 ASVCSSAGPN[illegible]FPGGVVSPQPHSSYYSGLSGPQHPFYNRAAVPATGSSLL---LSSMTV 1156
Db       1144 QAYPPYQMPQNEYPAN ---ELRSRNLSTSTEPVTPLINSPS-DSFGDDILQTKLTDLTV 1198

Qy       1157 DVVCEK[illegible]QIEGLLQNMMPQYCTTIKKANINGRVLSQCNIDELKKEMAMNFGDWHLFRSM 1216
Db       1199 EGVISL[illegible]PIEDM KFALPKLAPVLRENAINGPVLKHCDMFDLKSVLGLSFGHWELFRLL 1257

Qy       1217 VLEMRS[illegible]  SQVVPEDRRFLNENS[illegible]APV --- -PHGESAPRG---SHTELP 1258
Db       1258 ITTLRL[illegible]RKQLRQQQPGALEAPSNVPMIKDVTDALMQPPRRSLSRKNSVSHMEKQ 1317

Qy       1259 IT-- ELS[illegible]FYTLNFSFEELNTLGLDEGAPHHSNLSWQSQTRRTPSLSSL-------N 1308
Db       1318 VTLEEQM[illegible]FLQTLNEEAYE-DVASSERFSPTGEMLAAVAQLQLAPIRESSEFGSPSDD 1376

Qy       1309 SQDSSIEIRK[illegible]DKVQ---AEY-RDAYREYIAQMSQL------EGGTGSSTI---SGRSS 1355
Db       1377 QKQYGVKIS[illegible]NNNNQYLHAEYNRSVSSHSLQSLSTLVGAPVGHGGSGGSHLHLGNGNDL 1436

Qy       1356 PHSTYYI[illegible]S------GGSIHSTLEQ 1377
Db       1437 SDSTLDLMHVLSVFGGGGGGGYHRASRQ 1464

RESULT 13
AAE25146
ID   AAE25146 standard; Protein; 1498 AA.
XX
AC   AAE25146;
XX
DT   30-OCT-2002  (first entry)
XX
DE   Fruit fly ARMS protein.
XX
```

```
KW   Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
KW   ephrin; receptor tyrosine kinase; CNS; cancer; neuronal; cell;
KW   diagnostic; imaging; fruit fly.
XX
OS   Drosophila sp.
XX
FH   Key             Location/Qualifiers
FT   Region          376..415
FT                   /note= "N-terminal region"
FT   Region          594..697
FT                   /note= "SAM 1; between SAM 2 and TM"
FT   Region          740..884
FT                   /note= "Carboxy tail"
FT   Region          980..1192
FT                   /note= "Carboxy tail"
FT   Domain          1193..1255
FT                   /note= "SAM domain"
FT   Binding-site    1496..1498
FT                   /note= "PDZ binding site"
XX
PN   WO200250273-A2.
XX
PD   27-JUN-2002.
XX
PF   19-DEC-2001; 2001WO-US049445.
XX
PR   21-DEC-2000; 2000US-0258399P.
XX
PA   (UYNY ) UNIV NEW YORK STATE.
XX
PI   Chao MV, Kong H;
XX
DR   WPI; 2002-509800/54.
DR   N-PSDB; AAQ41538.
XX
PT   Isolated ankyrin repeat rich membrane spanning (ARMS) polypeptide that
PT   is a target for phosphorylation by neurotrophin receptor
PT   tyrosine kinases, useful as a marker for cancer
XX
PS   Disclosure; Page 1; 1[illegible] pages; English.
XX
CC   The invention relates to ankyrin repeat rich membrane spanning (ARMS)
CC   protein which is a down stream target of neurotrophin and ephrin
CC   receptor tyrosine kinases. ARMS [illegible] useful as a marker for cancer
CC   growth cone [illegible] ARMS [illegible]
CC   the biological activity of [illegible]
CC   for neuronal cells which [illegible]
CC   synaptic channels [illegible]
CC   distribution of ARMS [illegible]
CC   growth cones. ARMS DNA is useful [illegible]
CC   The present sequence is from [illegible]
XX
SQ   Sequence  1498 AA;

  Query Match          26.8%;  Score [illegible];  DB [illegible];  Length 1498;
  Best Local Similarity   57.6%;  Pred. No. [illegible];
  Matches  552;  Conservative  272;  Mismatches  [illegible];  Indels  [illegible];  Gaps  36;

Qy          1 MSVLISQSVINYVEEENIPACRAI[illegible]                    59
              |  |:::  ::  :  |:|:  :  :  :  |  :  :  |  :  |:  :|
Db         23 MDSLGHRALLQYIENNDISERAI[illegible]                    82

Qy         60 NGANCNLEDLENWTALHSASKEGH[illegible]                   119
              ||:  |||:: ||| :|:  :  :  |  :  |  :  |  :  |:  :
Db         83 RGADVQAEDLLNWTALDCASENGH[illegible]                   142

Qy        120 VELLLSHGANPSVTGLQYSVYPI[illegible]                     179
              | |||  |  :  |  :  :  :  |  :  |  :  |  :  |:  :|
Db        143 VRLLLDKGADGNAHG-NYHLGAL[illegible]                     201

Qy        180 RKGHLECVKHLLANGADVTQEGA[illegible]                     239
              |:  :|   ||  |:|  |  :|  :  :  :  |  :  :|  :  |  :
Db        202 RRGNVEIVDTLLKAGANVNAPDM[illegible]                     261

Qy        240 ALMIAPKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIR 299
              |:  |||:  :  :|  :  |:  ||  :||  |  :|  :  :|  :  :|  :  :||:
Db        262 ALCIAPREGFQDIAASLIAAGAYINIQDRGADTPLIHAVKAGHRTVVEALLKKHADVDIQ 321

Qy        300 GQDNKTALYWAVEKGNATMVRDLLQKNPDTEICTKDGETPLIKATKMRNIEVVELLLDKG 359
              |:| :||:  |||  :  :  :  :  |||  :  ||  :|||:  :  :|:|:|  :|:|:
Db        322 GKDRKTAIYTAVEKGHTPIVKLLLAINPDLESATKDGDTPLDRAVRNRNLEIVHLLLDRK 381

Qy        360 AKVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSI 419
              ||:  :  :  |:  :|  :  :  :  ||||  :|:|:  ||:||||:  :  ||:|:
Db        382 AKVTAIDFDGTCLHIAMRARSKTIVEALLRNPKHSQLLYRANKAGETPYNIDSLHQKTI 441

Qy        420 LTQIPGANPLSPTETDGDMLGYELYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKK 479
              |  :||  :  :  :  ,|||:  |||:  :|:|:  :|  |||||:|||||.:,|  |
Db        442 LGQVPDAPSLNTNEDSEGMLGYELYSSALADVLSEPTLTTPITVGLYAKWGSGKSFLLNK 501

Qy        480 LEDEMKTPAIQQTEPLPQFSWLIVELTLLLCGGLGLVFAFPV-DTNLAIAISLSFLALIY 538
              | |:  :  |  :  :  |:  :|:  :|:  :  :  :  :|  :  :  :  |  ,
Db        502 LRDEMNNPAPQWASPPIRTSGLLPIVGLHVAILIGTIVGLSTWSAVVGVSAAVGFLLLAY 561

Qy        539 IFFIVIYRSRREGESWNWAWALSTPLARHIGYLELLFKLMFVNPPELPEQTTKALPVRF 598
              :  :  :  ||:::  :  :  |:  :|  :|  :|  :  :  ||||
Db        562 LLLAAVKS-NYQ--MDMQWAYSVCHGLEKRMTSLRLILQVAFCHPPG-PQSDSQAKPVRF 618

Qy        599 LFTDYLRI  VGGETSLAEMIATLSDACEREPGPLATRLFRVFRTE--ESQGKKKWKKTC 656
              . :  |  :  :: ::  |:|  ||  :|: |:||:| |: :  :  :  |
Db        619 HFAEANSA  T-GDGAVAHMLAALLDAIESHYGWLATRLYRAFRPKCLKVDVGWRWRRMC 677

Qy        657 CLPSFVIPLFIVGCIIAGITLLAIF ----RVDPKHLTVNAILISIASVVGLAFVLNCRT 711
              |:|  :|:  :  :: |:'  :  :  :|:  :  |  '|:|:'  :
Db        678 CIPIVIIPLALVTVVTGISLTVAYFIFADEKEKEHILV--ALYVIAAVMGTLICTHLHV 735

Qy        712 WWQVLRCFLNQPRKRLHSAASKLHKLKSFGFMKVLKCEVELMARMAKTIDSFTQNQTRLV 771
              :|  |  :  |  :  :  ||  :  |  :|:|  :  :|  :  :|:||  :'|'|
Db        736 LAKVFINFFTSHIRVLKRAV----RSFESAPLTNLGAEVAVMTDMVKCLDAFTNQQSRLV 791

Qy        772 VIIDGLDKVSDKVLQMLDTVRVLFS-KQPFIAIFASDPHIIIKAINQNINSVLRDSNI 829
              :  '  :::' :: |:  '  :  :  :'|||:  |  :  :  :
Db        792 SVIDALS-TERITTLLNAVQTLLSNFNPFFVLLISVDPHVIAKAAEANSRRLFTFSGI 851

Qy        830 NGHDYMPNIVHLPVFINSRGLSNAPKFIVTS-  ATNGDITCND--TTGTQEDTDRRV 886
              ||:  :  :|:  :  :  :  :|:  :  |  :  |  |  :||:
Db        852 GGHDFSS'HLPVYLQNSGIPEVQRALMTALK-KRSGGGDYQTDDGDTLG-HSVSARRL 910

Qy        887 SLNS  KI  APTAAA  RDTYPRRQMQPTIIRQMSFDIIKL 926
              :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        911 SNASL  LKSLSRIAPGSSGKKLRTKFSVAIRGSNLERLQQNPQTV  LPGLSRI 965

Qy        927 LVTEDME  APQTMRKIQNIVSVTGPIQPANLITPNWDRLASWINLTEQWPYRTSWDIL 986
              :::'|'|  :|:  |':|  :  :  |':|  |:  :  |:|  |  ||||:|  '  |  :::
Db        966 VLTDDYP  VNPRSVRRLMNVIYITVRLLKAPQIEFSWYRLSSWINLTEQWPLRASMIVL 1025

Qy        987 YLEE  TE  FDQMTLKTMYERISKNIPTTKDVEPLLEIDGDIRNFEVFLSSRTPVLVAR 1043
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       1026 HHDQFVLQAIESVSIQSVYEKLRPKLAYLREAAPLLELDRDERKLDAFLQLHKSDLLVA 1085

Qy       1044 DVKTFLE  INLDPKLREIIADVRAAREQINIRGLAYPPLPLHEGP----PRFPSGYSQP 1099
              :  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       1086 DLRIFLP  INLDPYLRKVLKEDQQTIE--DESSLVIQARPSVSNTMRQFPAPTTYVPSP 1143

Qy       1100 ASVCSSASPNGPFPGGVVSPQPHSSYYSGLSGPQHPFYNRAAVPATGSSLL---LSSMTV 1156
              :  :  |  :|  :  :  |  |  :  :  |  |  :  :  |  :'  |:  :'
Db       1144 QAYPPYQVPLNEYPAN----ELRSRNLSTSTEPVTPLINSPS-DSFGDDILQTKLTDLTV 1198

Qy       1157 DVVCEPLA  EGLDQNMMPQYCTTIKKANTNGRVLSQCNIDELKKEMAMNFGDWHLFRSM 1216
              :  |  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       1199 EGVISLLS  EDM-KPALPFKLAPVLRENANGRVLKHCDMPDLKSVLGLSFGHWELFRLL 1257

Qy       1217 VLEMRSVP  - SQVVPEDPRFLNENPSAPV  --------PHGESARRS---SHTELP 1268
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       1258 ITTLRSTP  RRKQQPGQQQPGALEAPSNVPM  KDVTDALMQPPRESLSRKNSVSHMEKQ 1317
```

```
Qy       1259 LT --ELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSL    ----N 1308
              ||   ::    :  ||     | :    :  :|    |:  :| :  |           :
Db       1318 VTLEEQMICGTLQTLNEEAYE DVASSERFSPTSEMLAAVAQLQLAPIRESSEFGSPGDD 1376

Qy       1309 SQDSSIEISKLTDKVQ---AEY-RDAYREYIAQMSQL  -    EGGTGSSTI- -SGRSS 1355
              :   ::|,   :  |   |||  '     :  :  !          ::|   :  :
Db       1377 QKQYGVKISNNNNNNQYLHAEYNRDVSSHSLQSLSTLVGAIVGHDSGSHHLHLDNGNDL 1436

Qy       1356 PHSTYYIGQSSS-- -- GGSIHSTLEQ 1427
              |  :  |     :| '   .
Db       1437 SDSTLDLMHVDSVFSGGGGGGYHRAHS 1464


RESULT 14
AAE03645
ID   AAE03645 standard; Protein; 513 AA.
XX
AC   AAE03645;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Human extracellular matrix and cell adhesion molecule-9 (XMAD-9).
XX
KW   Human; extracellular matrix and cell adhesion molecule; XMAD;
KW   gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
KW   Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
KW   sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
KW   inflammatory disorder; acquired immune deficiency syndrome; AIDS;
KW   Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
KW   Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
KW   glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
KW   osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
KW   infection; cell proliferative disorder; actinic keratosis; myeloma;
KW   arteriosclerosis; nootropic; antihypovalsant; antihypotoid; nephrotropic;
KW   neuroprotective; dermatological.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          94..108
FT                   /note= "Adrenocorticotrophin receptor"
FT   Domain          196..198
FT                   /note= "Rgd cell interaction motif"
FT   Region          328..335
FT                   /note= "Beta integrin family"
XX
PN   WO200142285-A2
XX
PD   14-JUN-2001.
XX
PF   05-DEC-2000; 2000WO-US32999
XX
PR   10-DEC-1999;   98US-0172852.
PR   16-DEC-1999;   99US-0172354.
XX
PA   (INCY ) INCYTE GENOMICS INC.
XX
PI   Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;
PI   Baughn MR, Lu DAM, Shah P, Au Yeung C;
XX
DR   WPI; 2001-381632/40.
DR   N-PSDB; AAD08053.
XX
PT   New human extracellular matrix and cell adhesion molecules and
PT   polynucleotide sequences encoding them, useful for diagnosis,
PT   prevention, treatment of genetic, autoimmune and cell proliferative
PT   disorders -
XX
PS   Claim 1; Page 101-102; 135pp; English.
XX
CC   The present sequence is a human extracellular matrix and cell
CC   adhesion molecule (XMAD). The XMAD is used for screening a compound for
CC   effectiveness as an agonist or antagonist of XMAD. The identified agonist
CC   or antagonist are used for treating a disease or condition associated
CC   with decreased or increased expression of functional XMAD. The
CC   polynucleotides encoding XMAD are useful in somatic or germline gene
CC   therapy to correct a genetic deficiency, to express a conditionally
CC   lethal gene product and to express a protein which affords protection
CC   against intracellular parasites and also for diagnosis of disorders
CC   associated with expression of XMAD. They are also used for generating
CC   hybridisation probes useful in mapping the naturally occurring genomic
CC   sequences and to create knock in humanised animals (pigs) or transgenic
CC   animals (mice or rats) to model human diseases. Oligonucleotide or longer
CC   fragments derived from the polynucleotide sequences may be used as
CC   elements on a microarray. Antibodies which specifically bind XMAD may be
CC   used for the diagnosis of disorders associated with the expression of
CC   XMAD, or in assays to monitor patients being treated with XMAD. Diseases
CC   diagnosed, prevented or treated include genetic disorders such as
CC   adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
CC   disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
CC   autoimmune/inflammatory disorders such as acquired immune deficiency
CC   syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
CC   atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
CC   glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
CC   osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC   bacterial, fungal, parasitic, protozoal and helminthic infections and
CC   cell proliferative disorders such as actinic keratosis, arteriosclerosis
CC   and cancer including breast, bladder, bone marrow, brain and uterus
CC   cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
XX
SQ   Sequence   513 AA,

  Query Match             25.5%;  Score 2260;  DB 22;  Length 513;
  Best Local Similarity   86.6%;  Pred. No. 4.4e-169;
  Matches  445;  Conservative   21;  Mismatches   46;  Indels    2;  Gaps    2;

Qy       1203 MAYNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELFLTEL 1262
              |  : |||  . |||  |||||: || |  |||||:|:||| | |||| |||:|| ' ' | |||
Db          1 MNYNFGDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 60

Qy       1263 SSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDK 1322
              |||:|||| ||||:|||| :|||:|||||:|||||:|||||||:|||||'|:'|||
Db         61 SSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDK 120

Qy       1323 VQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSFHSTYYIGQSSSGGSIHSTLEQERGKE 1382
              |||:|:|  :|||:|||:'  ||:  |||| :|||||:|||||||:|| || |:|:
Db         121 VQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSFHSTYYMGQSSSGGSIHSNLEQEKGKD 180

Qy       1383 GELKQEYKFMFLMKPGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSE 1442
              |   :|'  ||| . ||| . |:| . |:|  :|||:| |||'| | |||||': ||
Db         181 SEPKPDYKFMFLMKPGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE 240

Qy       1443 RPSLFQTDLKLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIRTKRLKAKQRECASPQE 1502
                ||  ' ||   | | :'|||'|.|||:       :,   |  : :: : ,
Db         241 RSSLFQTDLKLKGSGLRYQKLPSDEDESGTEESDNTP-LLKDDKDRKAEGKVERVPKSPE 299

Qy       1503 HSAEPIRTPIFAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADDSQLEKANLIEL 1562
              '|||'' |||  ||| :|||'|||||.:|||:|||||'|||  |||' |||||||
Db         300 HSAEPIRTPIFAKEYLSDALLDKKDSSDSGVRSSESSPNHSLHNEVADDSHLEKANLIEL 359

Qy       1563 EDEGHSGKPGPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLN 1622
              |: ' |||  :|||||'|||:.||||'.|||||'||||:|||||:|:|||||'|||||
Db         360 EDDSHSGKPGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLN 419

Qy       1623 RTPSTVTLNNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQNENLKSMAHKRS 1682.
              |||'':||' :|| :|||:|||:'||||||.|||  ||||, :||||||| |||
Db         420 RTPSTVTLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRS 479

Qy       1683 QRSSYTPLSKDASELH-AASSESTGFGEERESIL 1715
              '|||  |||   ||: ||| ||||| ||||.
Db         480 QRSSYTPLSKDPPELHAAASSESTGFGEERESIL 513


RESULT 15
ABG08697
ID   ABG08697 standard; Protein; 624 AA.
```

XX
AC ABG08697;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #8784
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US09841.
XX
PR 31-MAR-2000; 2000US-0540217
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639762/73.
DR N-PSDB; AAS72684.
XX
PT New isolated polynucleotide and encoded polypeptides useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 19056; 1499pp; English
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, diagnostics, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its fragments are also useful in medical
CC imaging of sites expressing (II). [illegible] are useful in the treatment of
CC disorders involving abnormal protein [illegible]
CC The polypeptide and polynucleotide [illegible]
CC diagnostics, forensics, gene mapping, [illegible]
CC responsible for genetic disorders [illegible] biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG08618 ABG08697 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 624 AA;

Query Match 24.3%; Score [illegible]; DB 22; Length 624;
Best Local Similarity 80.6%; Pred. No. [illegible];
Matches 435; Conservative 71; Mismatches 66; Indels 8; Gaps 7;

Qy 1183 KANINGRVLSQCNIDELKKEMAMNFSDWFERMVLEMKSVDVQVVPEDHRFLNENSSAP 1242
Db 86 EANINGRVLAQCNIDELKKEMNMNFSDWFERMTVLEMRNAESHVVPEDHRFLSESSSGP 145

Qy 1243 VPHGESARRSSHTELPLTELSSQTPYTLNPSPEELNTLGLIEGAPRHSNLSWQSQTRRTP 1302
Db 146 APHGEPARRASHNELPHTELSSQTPYTLNPSPEELNTLGLDEGAPRHSNLSWQSQTRRTP 205

Qy 1303 SLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYM 1362
Db 206 SLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYM 265

Qy 1363 GQSSSGGSSHHNTLEQERGKEGELKQEDGRKSFLMKRGDVIDYSSSGVSTNEASPLDPITE 1422
Db 266 GQSSSGGSSHHNLEQEKGKDSEPKPDDGRKSFLMKRGDVIDYSSSGVSTNDASPLDPITE 325

Qy 1423 EDEKSIQGSFLLPGKKSSERPSLFQTDLKLKGGSLRYQKLPSDEDESGTGRVQITPHCS 1482
Db 326 EDEKSIQGSFLLPGKKSSERSSLFQASLKLKGPGLRLSKLPSEEDECGAEESDNTP-LL 384

Qy 1483 KMIRTKFLKAHQRECASPQEHSAEPIRTFIKAKEYLSDALLDKK DSSDS-GVRSNESSP 1540
Db 385 KDDKDRKFGGFVERVPKSPEHSVEPIRTFIKAKEYLSDALLDNKGNSSDSWSVDPRESSP 444

Qy 1541 NH-SLHNEAATG-SQLEKANLIELEDEGHSGKRGMPHSLSGLQDPIIARMSICSEDKKSP 1598
Db 445 NHVSAQLYAADLPDLKRHNLIELEDDSHSGKRGIPHSLSGLQDPIIARMSICSEDKKSP 504

Qy 1599 SECSLI--ASSPEESWPACQKAYNLNRTPSTVTLNNNTAPTNRANQNFDEIEGIRETSQV 1656
Db 505 SRIASLDKFDSPEENWPACQKAYNLNRTPSTVTLNNNSAPANRANQNFDEMEGIRETSQV 564

Qy 1657 ILRPGPSFNFTAVQNENLKSMAHKRSQRSSYTRLSKDASELHAA-SSESTQFGEERESIL 1715
Db 565 ILRPSSSFNFTTIQNENLKSMTHKRSQRSSYTRLSKESPELHAASSSESTGFGEERESIL 624

Search completed: November 21, 2003, 12:44:33
Job time : 60.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 12:[illegible], search time 2.5 seconds
without alignments
2871.130 Million cell updates/sec

Title: US 10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEEENIEA[illegible]EIHEEEEESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96164682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5073.5	57.3	1031	2	T43458	hypothetical protein DKFZp434F0621.1 - human
2	1848.5	20.9	1498	2	T43446	hypothetical protein DKFZp434A0131.1 - human
3	1024	11.6	246	2	S48660	hypothetical protein DKFZp434A0131.1 - human
4	596	6.7	1881	1	A02078	ankyrin 1, erythro
5	565	6.4	426	2	AB0141	hypothetical prote
6	553	6.2	1856	2	B35049	ankyrin 1, erythro
7	553	6.2	1880	2	A35049	ankyrin 1, erythro
8	552	6.2	1961	2	T42716	ankyrin 3, splice
9	551.5	6.2	1940	2	T42715	ankyrin 3, splice
10	551	6.2	1943	2	T42713	ankyrin 3, splice
11	550	6.2	3924	2	S37431	ankyrin 2, neuronal
12	537.5	6.0	1862	2	I49502	ankyrin - mouse
13	526	5.9	4377	2	A65575	ankyrin 3, long sp
14	516.5	5.8	1848	2	S37721	ankyrin, erythrocy
15	512.5	5.8	1765	2	T42714	ankyrin 3, splice
16	479	5.4	2039	2	T15943	ankyrin-related un
17	464.5	5.2	1549	2	T13940	ankyrin - fruit fl
18	429	4.8	2584	2	T24156	hypothetical prote
19	429	4.8	2606	2	T24157	hypothetical prote
20	414.5	4.7	1423	1	I37275	death-associated p
21	409.5	4.6	791	2	T42691	hypothetical prote
22	397	4.5	1062	2	T30255	inversin - mouse
23	384.5	4.3	557	2	T46507	hypothetical prote
24	382	4.3	1062	2	T14151	Inv protein - mous
25	378.5	4.3	323	2	B47169	ankyrin-like repea
26	361	4.1	397	2	T46445	hypothetical prote
27	356	4.0	679	2	B45771	2'-5A-dependent RNA
28	347.5	3.9	741	2	A45771	2'-5A-dependent RNA
29	348.5	3.8	991	2	T25417	hypothetical prote
30	335	3.8	1401	2	S21527	alpha-latrotoxin p
31	325	3.7	1188	2	T19552	hypothetical prote
32	322.5	3.6	441	2	S74626	erythroid ankyrin
33	321.5	3.6	934	1	H71274	probable ankyrin -
34	302	3.4	1958	2	D82654	ankyrin-like prote
35	301	3.4	2004	2	D88948	protein ZK1005.1 [
36	300.5	3.4	1411	2	S30355	alpha-latroinsecto
37	296.5	3.3	627	2	E86190	hypothetical prote
38	294	3.3	1384	2	T00253	gene Ankhzn protei
39	292	3.3	406	2	AC2508	hypothetical prote
40	288	3.3	1781	2	S48460	probable membrane
41	287	3.2	517	2	T48283	ankyrin-like prote
42	286.5	3.2	462	2	F84725	ankyrin-like prote
43	285	3.2	1121	2	S30385	G9a protein - huma
44	279.5	3.2	668	2	T16184	ankyrin repeat pro
45	279	3.2	443	2	T86212	hypothetical prote

ALIGNMENTS

RESULT 1
T43458
hypothetical protein DKFZp434F0621.1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43458
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z27517
A;Accession: T43458
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1031 <AAA>
A;Cross-references: EMBL:AL133620
A;Experimental source: adult testis; clone DKFZp434F0621
C;Genetics:
A;Note: DKFZp434F0621.1

Query Match 57.3%; Score 5073.5; DB 2; Length 1031;
Best Local Similarity 95.2%; Pred. No. 4.9e-293;
Matches 974; Conservative 23; Mismatches 13; Indels 13; Gaps 1;

Qy 1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNFCGQTPLMLAAEQGNVEIVKELLKN 60
Db . MSVLISQSVINYVEEENIPALKALLEKCKDVDERNFCGQTPLMLAAEQGNLEIVKELIKN 60

Qy 61 GANCNLEDKDMTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120
Db 61 GANCNLEDKDMNTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

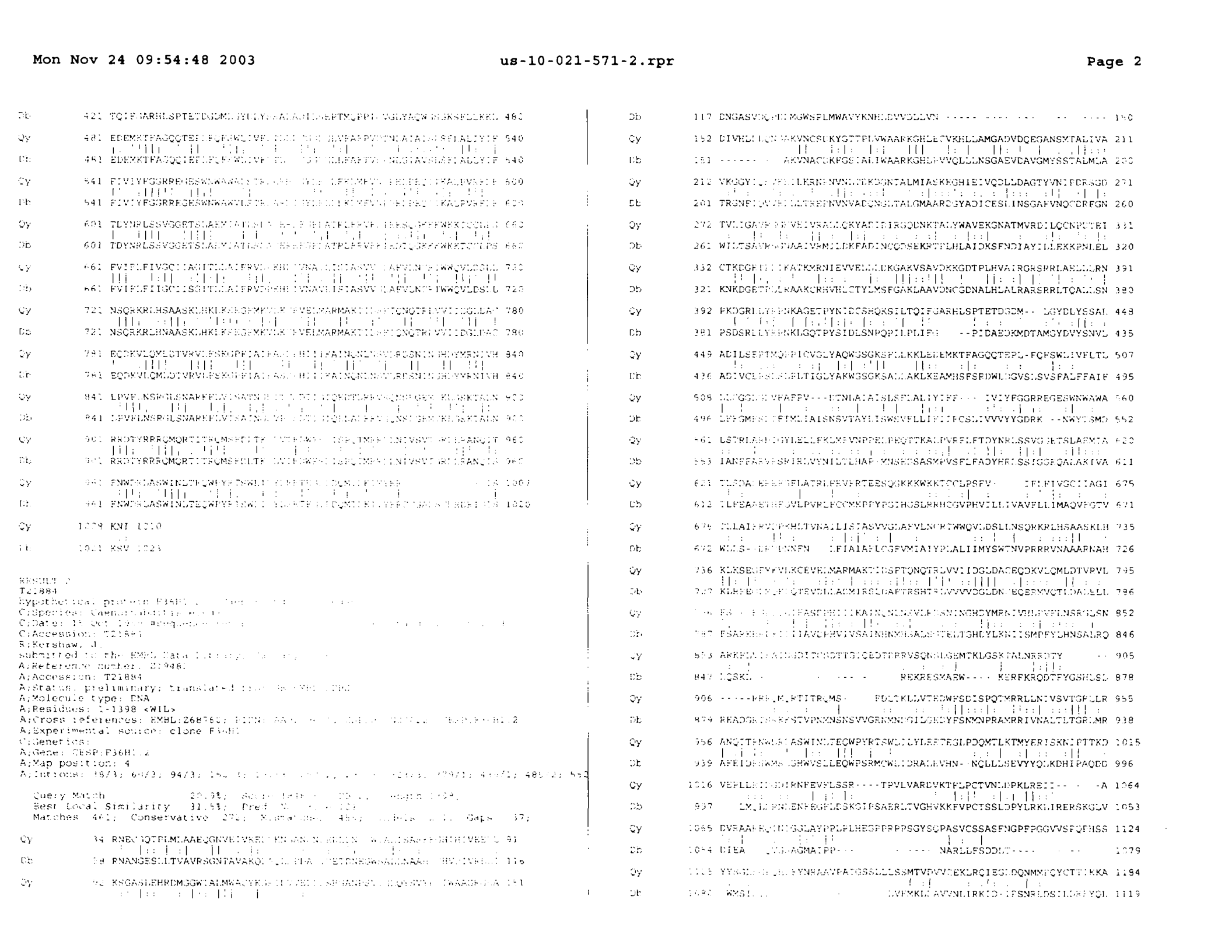
Qy 181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
Db 181 KGHLECVKHLIAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy 241 LMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
Db 241 LMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
Db 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy 361 KVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
Db 361 KVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy 421 TQIFGARHLSPIETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480



```
Qy       1186 NINGRVLSQCNIDELKKEMAMNFGKWHLERSMVIEMRSVEDVVPRDSPPFLNENSSAPVP 1244
Db       1120 NLQGLVLATCPLPELKDSMQLPLGLWTLIR-LLLST          ----- LKVEGSSP - 1162

Qy       1245 HGESARRSSHTELPLTELSNQTPYTLNF-PSLSNTLNLDESAIRHSNLSWSSQTRRTPSL 1304
Db       1163 -------PGLRVDVPRPALTTA-RDEEPH   IEEAAGAAIN-   SEKERAPLL 1202

Qy       1305 SSL ----------NSQSSSIELPLTKVQAEYRDAYPRTAQMSQLPSTG - -   1346
Db       1203 GSVRAEQRRRSTIVDNATELSIDH LN KLS-          NTLTEIP-DVNEMHFJH 1252

Qy       1347 -SSTISDRSSPHSTEYTDDNSAGTR TH                DTLRQAR RKEGELSQRD 1389
Db       1253 FSSSTDGPSPMADGFLPASVSAAPDVPP DNINDLERKASTADETQSRYTNSKENLL  ED 1310

Qy       1390 GRKSFLMKRGDVIDYSSSGVSTN 1411
Db       1311 ERSANPPAHVDLMRPDSDNTSSD 1333
```

```
RESULT 3
T46392
hypothetical protein DKFZp434E0620.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 11-Feb-2000
C;Accession: T46392
R;Ottenwaelder, B.; Obermaier, B.; Meyer, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46392
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-246 <AAA>
A;Cross-references: EMBL:AL136755
A;Experimental source: adult testis; clone DKFZp434E0620
C;Genetics:
A;Note: DKFZp434E0620.1

  Query Match             11.6%;  Score 1[illegible];  DB 2;  Length 246;
  Best Local Similarity   80.7%;  Pred. No. 1.[illegible]e-53;
  Matches  205;  Conservative   10;  Mismatches   19;  Indels   20;  Gaps    3;

Qy       1463 LPSDRPESGTGSVQITPHPNKMIP VRI-AA[illegible]  1512
Db         12 LKMCFIPRAENVVPVI                                  62

Qy       1513 LDKKLSSDSGVRSNARQHLQNLD                            1544
Db         63 LDKKISSDSGVRSSESSPMLSLNDLA                        112

Qy       1545 PIIARMSICSEDKKSPSECSLIASA EEIWPAT RAYNLNP  ITVTLMNNDAPANRANQ 1604
Db        113 PIIARMSICSEDKKSPSECSLIASA EEIWPAT RAYNLNP  ITVTLMNNDAPANRANQ 172

Qy       1605 NFDEIEGIRETSQVILRPGSPNP TAVCLANNLPDMAHP-QL  YEIPLSKIASELH-AAS 1661
Db        173 NFDEMEGIRETSQVILRPGSASPNT TICLPSNLPDMTHPASLA YEIPLSKITPELHAAAS 232

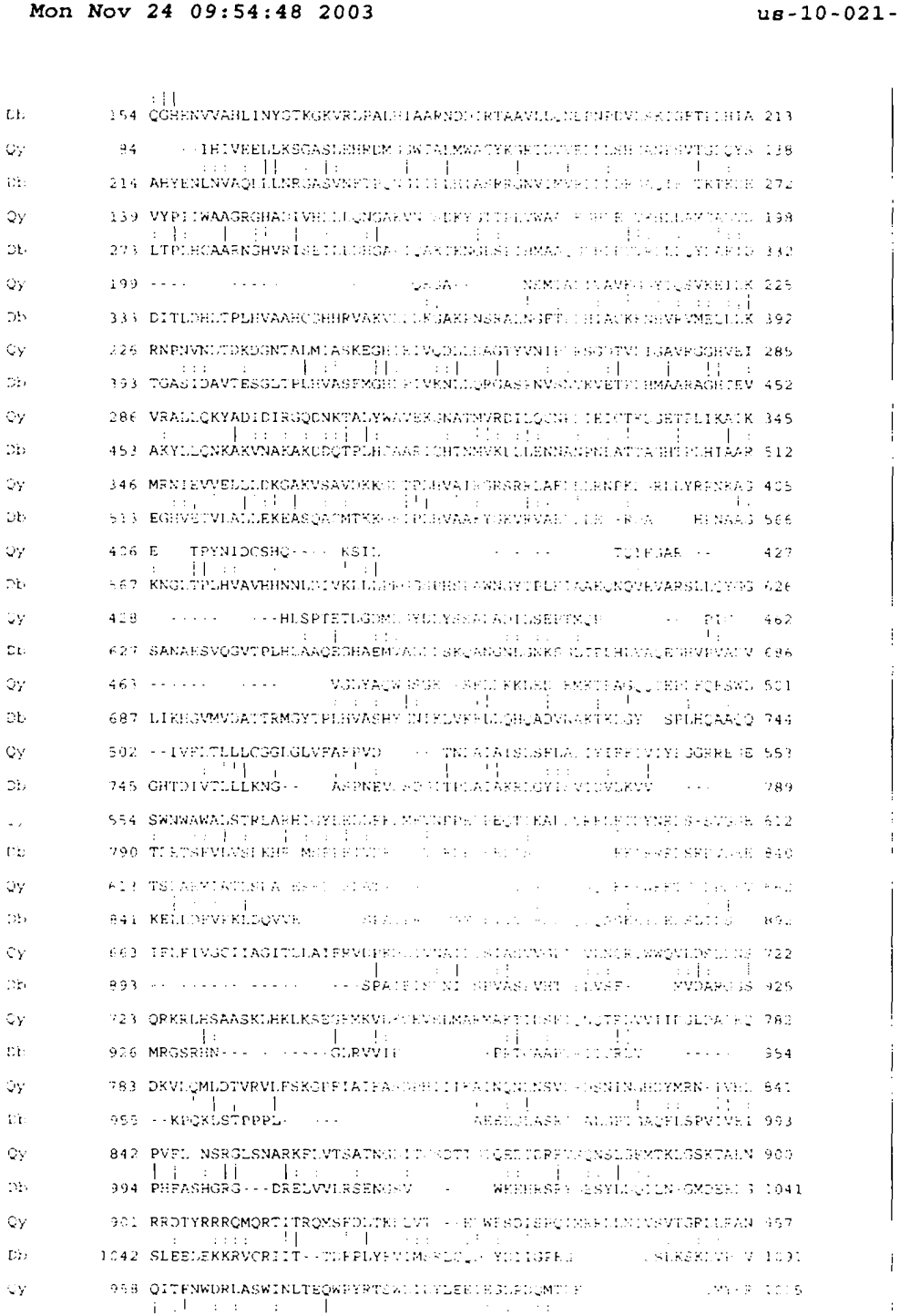
Qy       1662 SESTGPGEERESIL 1719
Db        233 SESTGPGEERESIL 246
```

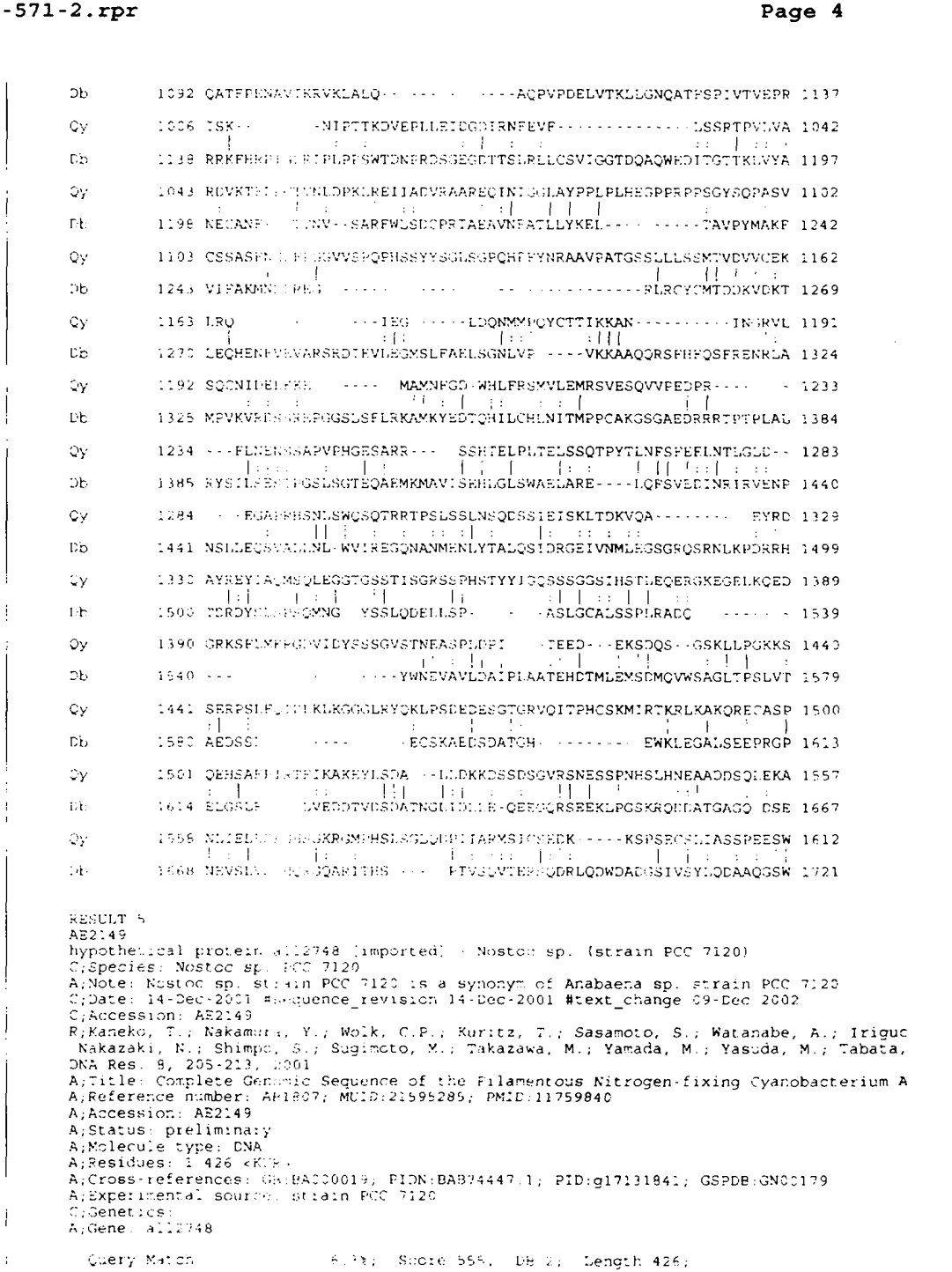
```
RESULT 4
SJHUJK
ankyrin 1, erythrocyte splice form 1 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin R
N;Contains: ankyrin 2.2
C;Species: Homo sapiens (man)
C;Date: 1[illegible] Sep 1991 #sequence_revision 30-Sep 1991 #text_change 09-Jun 1999
C;Accession: S08275; A33219; PC2220; A[illegible]
R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A;Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure
A;Reference number: S08275; MUID:90158830; PMID:2137557
A;Accession: S08275
A;Molecule type: mRNA
A;Residues: 1-1881 <LUX>
A;Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A;Accession: A33219
A;Molecule type: protein
A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753,828-833,'X',835-855,'
X',1367;1383-1427;1[illegible]1686-1698,'D',1700;1763-1772 <LUX>
A;Note: 845-Arg and 1[illegible] Thr were also found
R;Bermann, J.; Bare[illegible] M.; Prade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membra
A;Reference number: PC2220; MUID:95071348; PMID:7526850
A;Accession: PC2220
A;Molecule type: protein
A;Residues: 910-929 <BER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchange
A;Reference number: A[illegible]443; MUID:90285190; PMID:2141335
A;Accession: A35443
A;Molecule type: protein
A;Residues: 'X',5,'X',7-12;403-417,'X',419-422,'H',424,'LQ';797-800,'L',802-814;862-86
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2 8p11.2
C;Superfamily: ankyrin, ankyrin repeat homology
C;Keywords: alternative splicing; phosphoprotein
F;2-188/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F;2-1512,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>
F;2-827/Domain: 89K #status predicted <DOM1>
F;2-827/Region: anion exchange protein binding
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-175/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;[illegible]-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>
F;828-1382/Domain: 72K #status predicted <DOM2>
F;828-1382/Region: spectrin binding
F;1383-1881/Domain: 58K #status predicted <DOM3>

  Query Match              6.3%;  Score 556;  DB 1;  Length 1881;
  Best Local Similarity   20.6%;  Pred. No. 3.6e-24;
  Matches  395;  Conservative  270;  Mismatches  635;  Indels  620;  Gaps   75;

Qy         21 LKALLEKGVDVDERNEQGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWTALISASK 80
Db         94 VREILVKHANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQ 153

Qy         81 EGH                                                        83
```

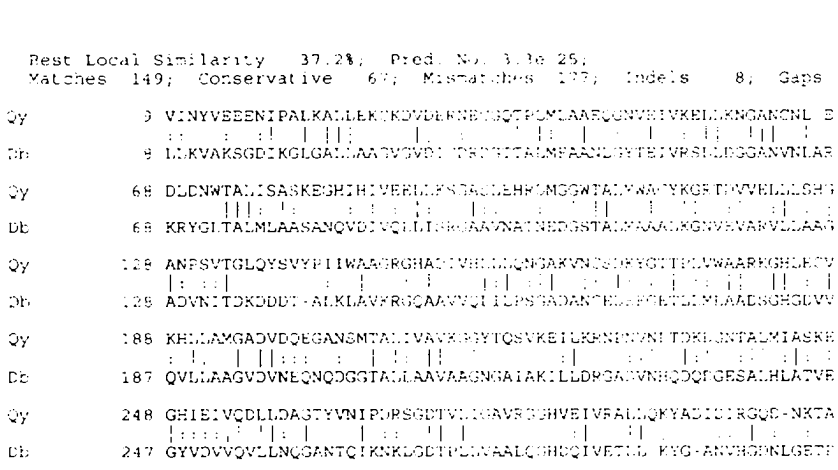




RESULT 5
AE2149
hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec 2002
C;Accession: AE2149
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A;Reference number: AR1807; MUID:21595285; PMID:11759840
A;Accession: AE2149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1 426 <KUF>
A;Cross-references: GB:BA000019; PIDN:BAB74447.1; PID:g17131841; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2748

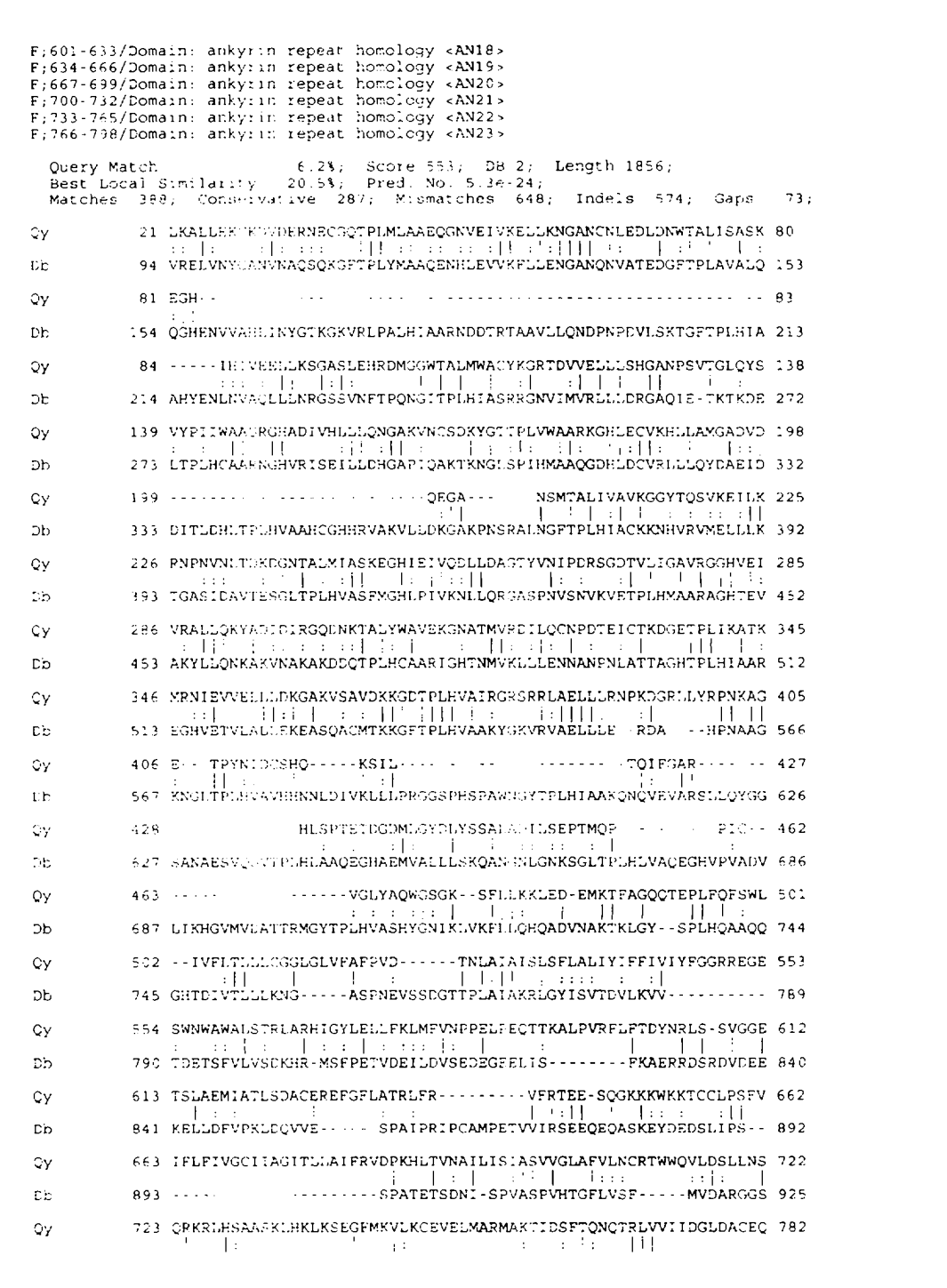
Query Match [illegible]; Score 555; DB 2; Length 426;

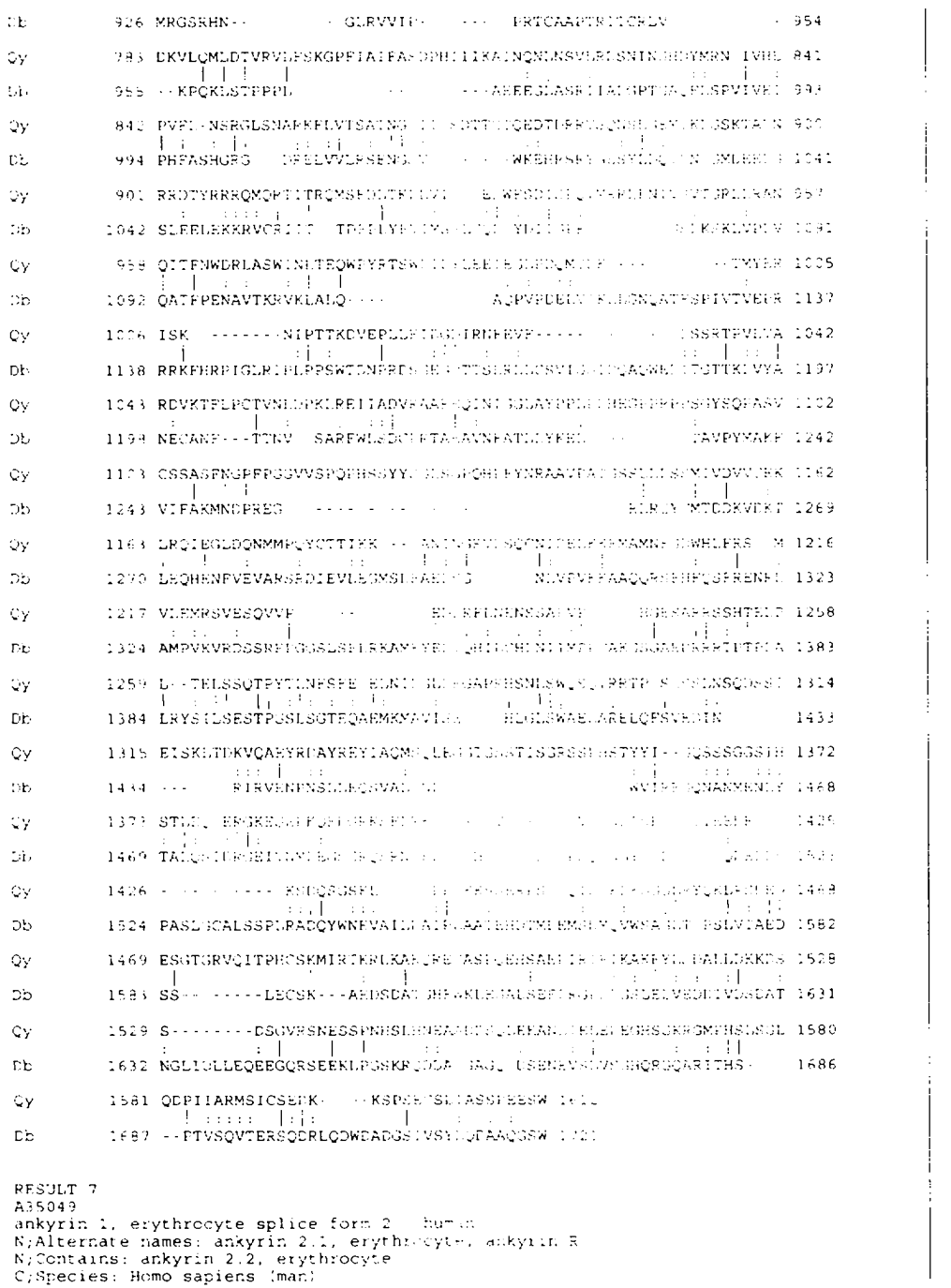
Best Local Similarity 37.2%; Pred. No. 3.3e-25;
Matches 149; Conservative 67; Mismatches 127; Indels 8; Gaps 6;



RESULT 6
B35049
ankyrin 1, erythrocyte splice form 3 - human
N;Alternate names: ankyrin 2.1, erythrocyte ankyrin B
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Oct 1998
C;Accession: B35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin
A;Reference number: A35049; MUID:90175470; PMID:1689449
A;Accession: B35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1856 <LAM>
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118777; OMIM:1...
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN1>
F;77-109/Domain: ankyrin repeat homology <AN2>
F;110-142/Domain: ankyrin repeat homology <AN3>
F;143-171/Domain: ankyrin repeat homology <AN4>
F;172-204/Domain: ankyrin repeat homology <AN5>
F;205-237/Domain: ankyrin repeat homology <AN6>
F;238-270/Domain: ankyrin repeat homology <AN7>
F;271-303/Domain: ankyrin repeat homology <AN8>
F;304-336/Domain: ankyrin repeat homology <AN9>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

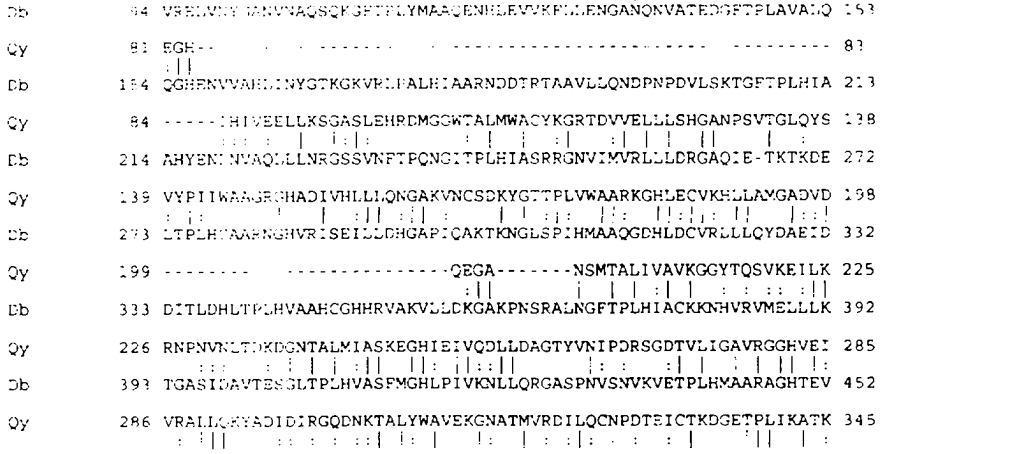
Query Match 6.2%; Score 553; DB 2; Length 1856;
Best Local Similarity 20.6%; Pred. No. 5.3e-24;
Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;

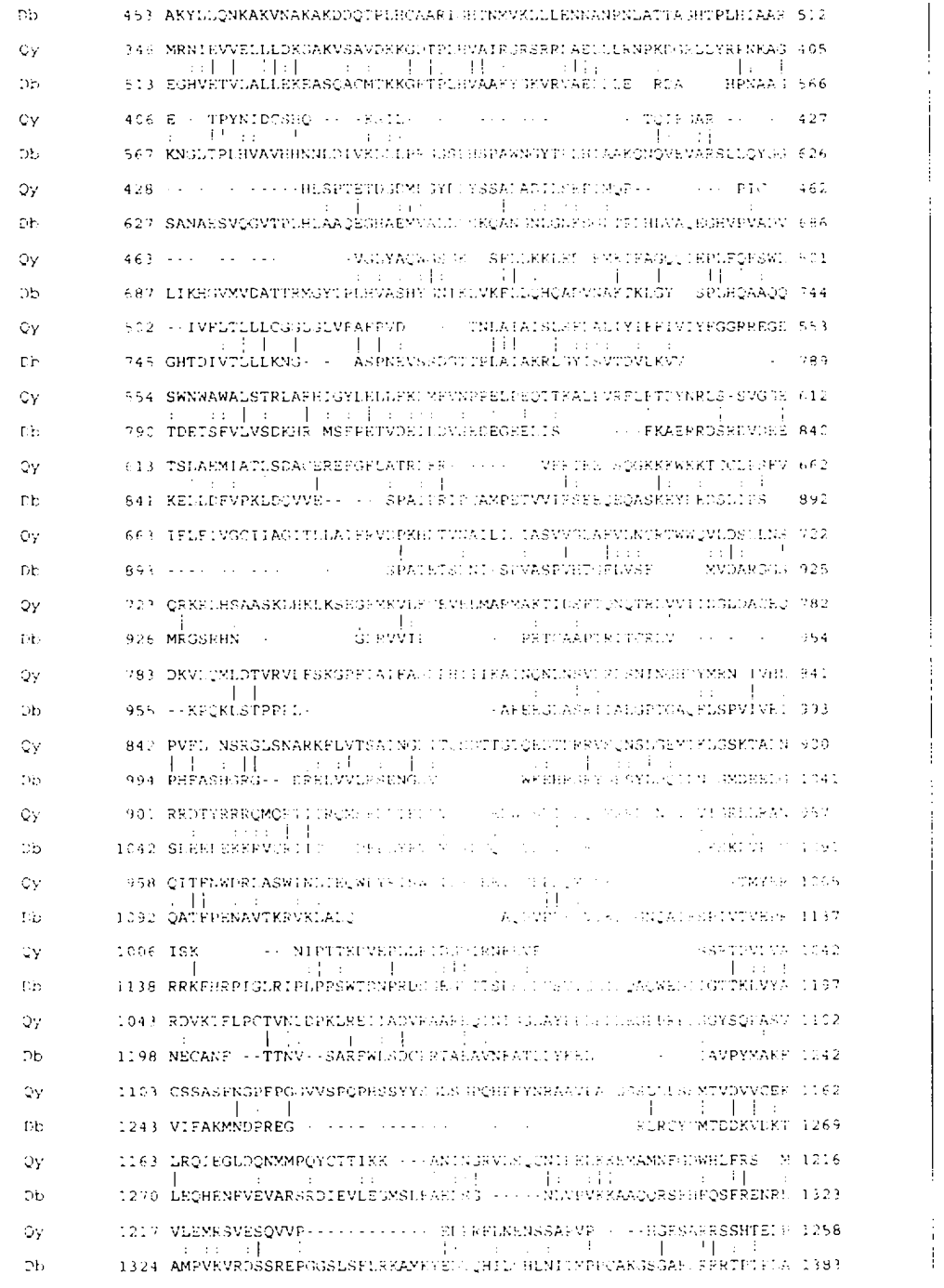


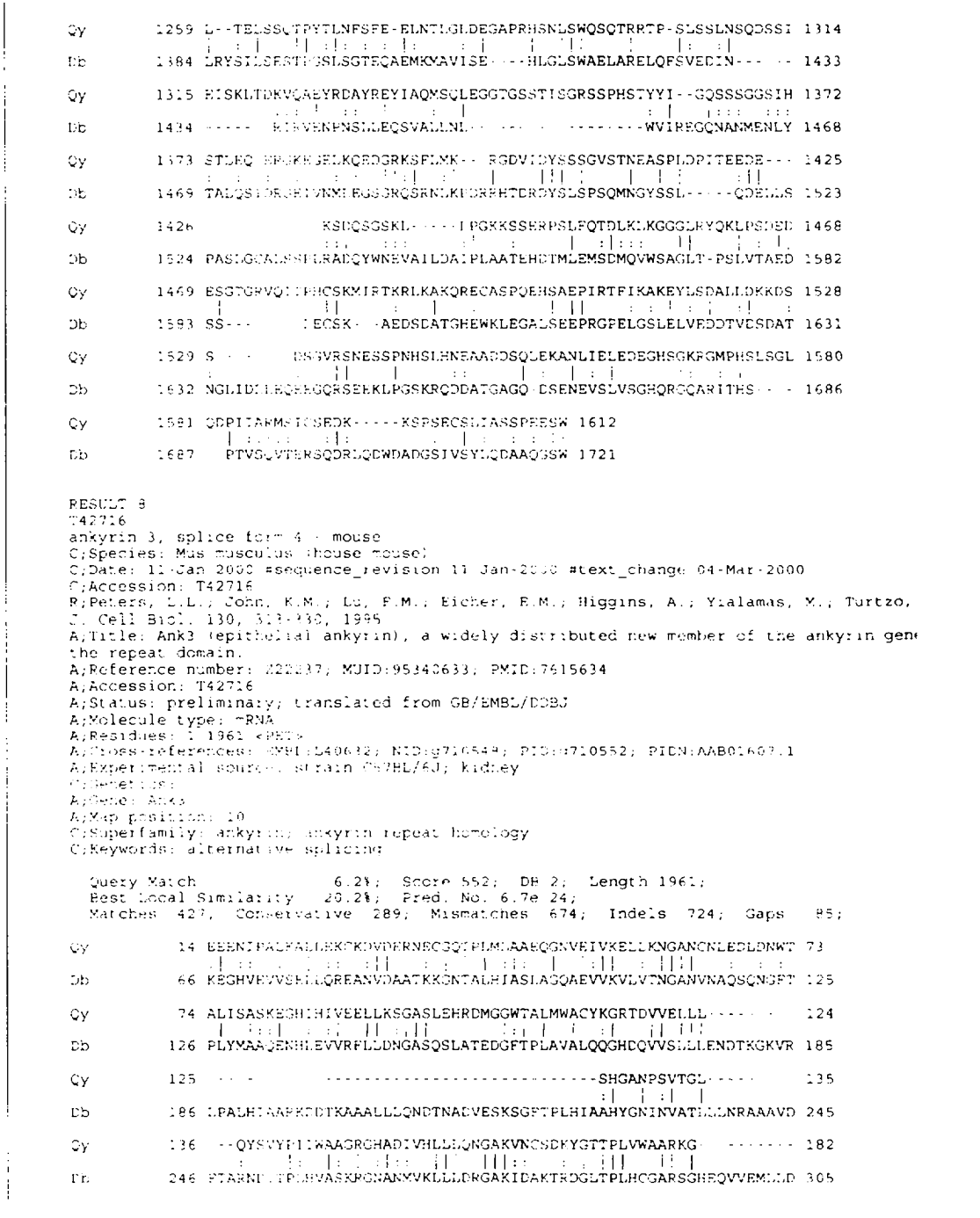


RESULT 7
A35049
ankyrin 1, erythrocyte splice form 2 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C;Accession: A35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A[illegible]; MUID:90175[illegible]; PMID:1689849
A;Accession: A35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1880 <LAM>
A;Cross-references: [illegible]
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:[illegible]; OMIM:[illegible]
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; cytoskeleton
F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 553; DB 2; Length 1880;
Best Local Similarity 20.5%; Pred. No. 5.4e-24;
Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;

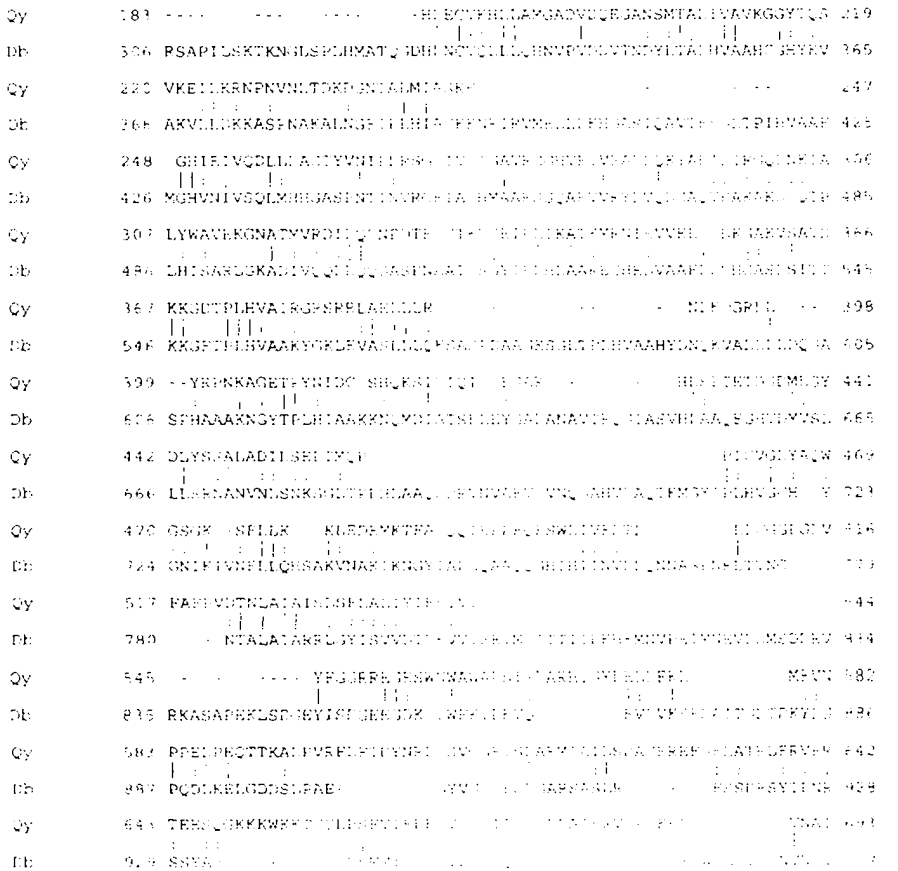


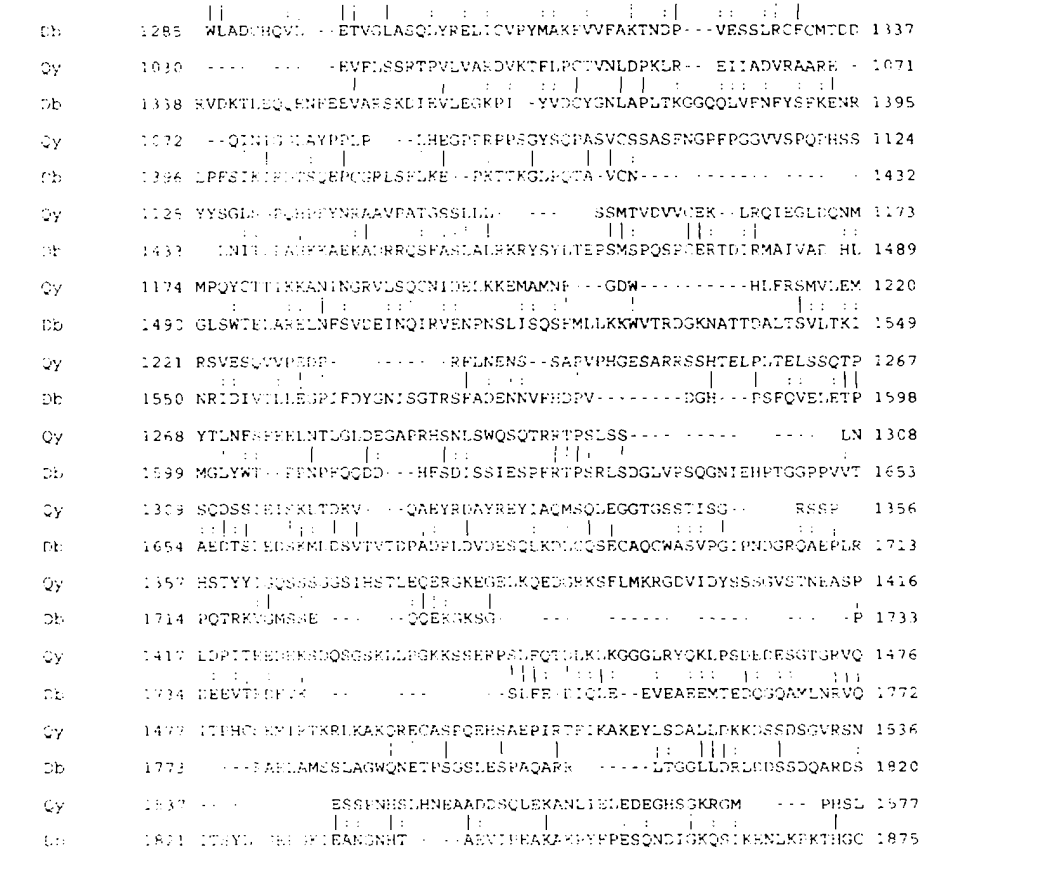




RESULT 8
T42716
ankyrin 3, splice form 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42716
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene the repeat domain.
A;Reference number: A22837; MUID:95340633; PMID:7615634
A;Accession: T42716
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1961 <PET>
A;Cross-references: GB:L40632; NID:g710549; PID:g710552; PIDN:AAB01607.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

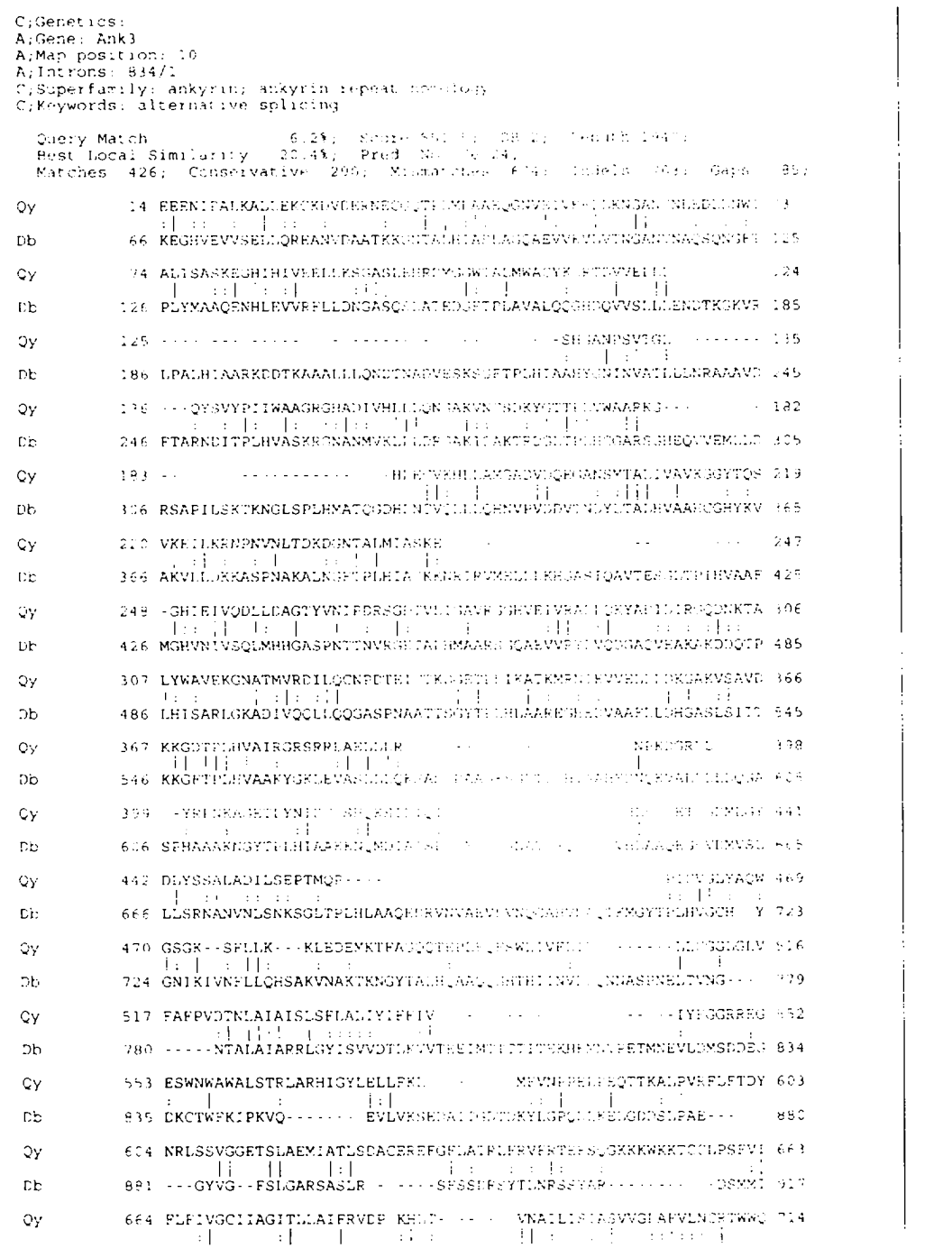
Query Match 6.2%; Score 552; DB 2; Length 1961;
Best Local Similarity 20.2%; Pred. No. 6.7e-24;
Matches 427; Conservative 289; Mismatches 674; Indels 724; Gaps 85;

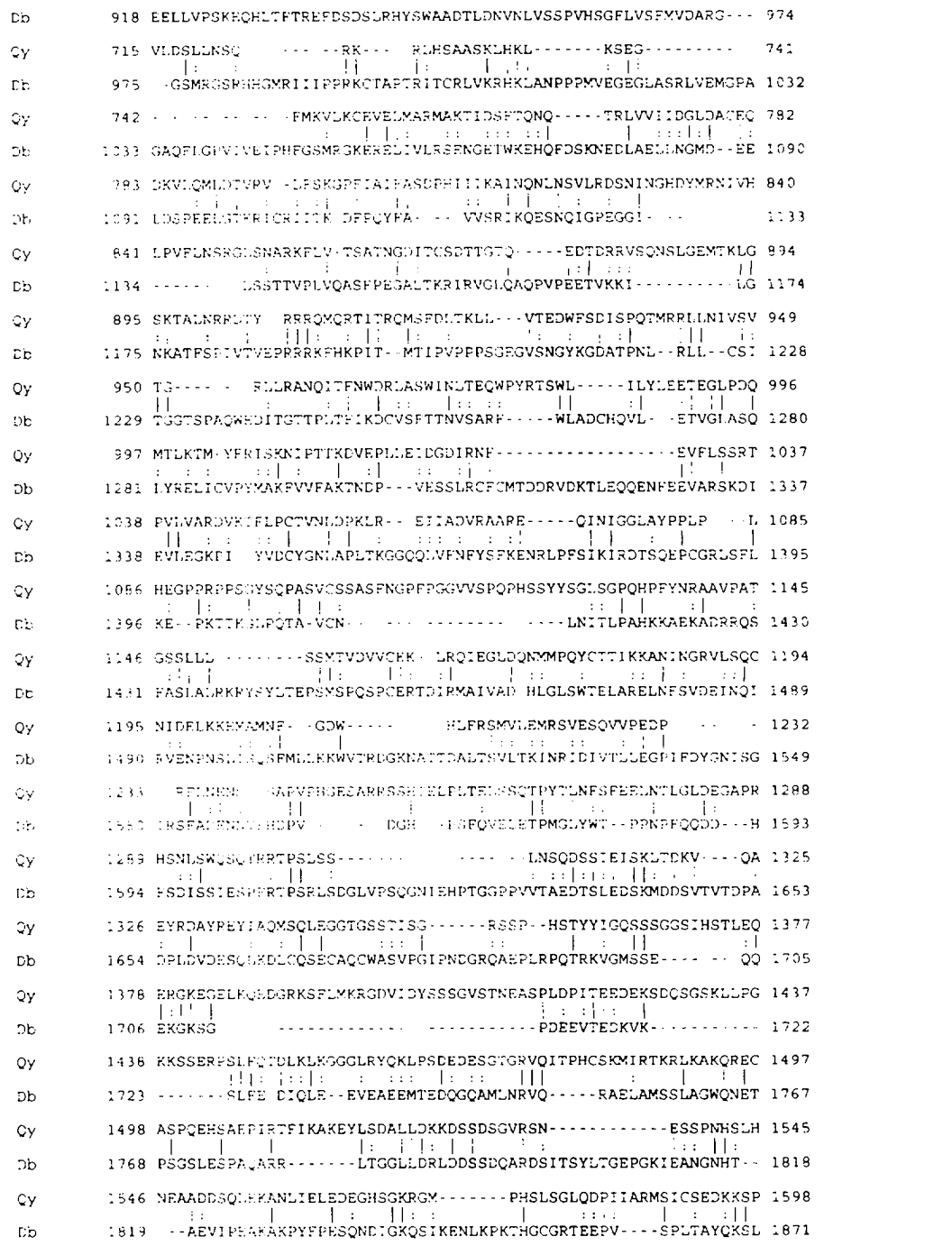


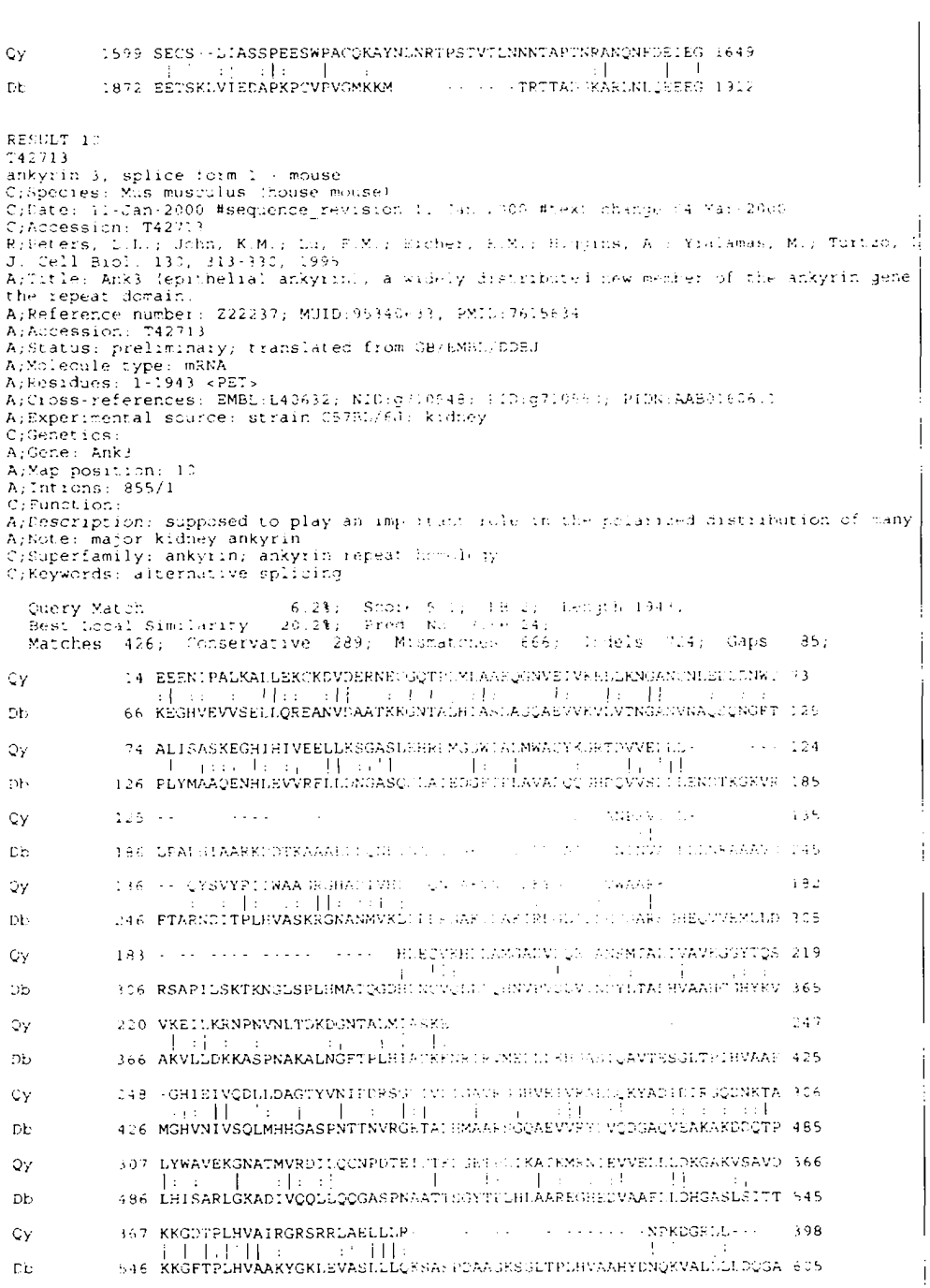


RESULT 9
T42715
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan 2000 #sequence_revision 11 Jan-2000 #text_change 04-Mar-2000
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gen the repeat domain.
A;Reference number: Z27137; MUID:95340633; PMID:7615634
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PET>
A;Cross-references: EMBL:U40632; NID:g710548; PID:g710549; PIDN:AAB01604.1
A;Experimental source: strain C57BL/6J; kidney

C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 834/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

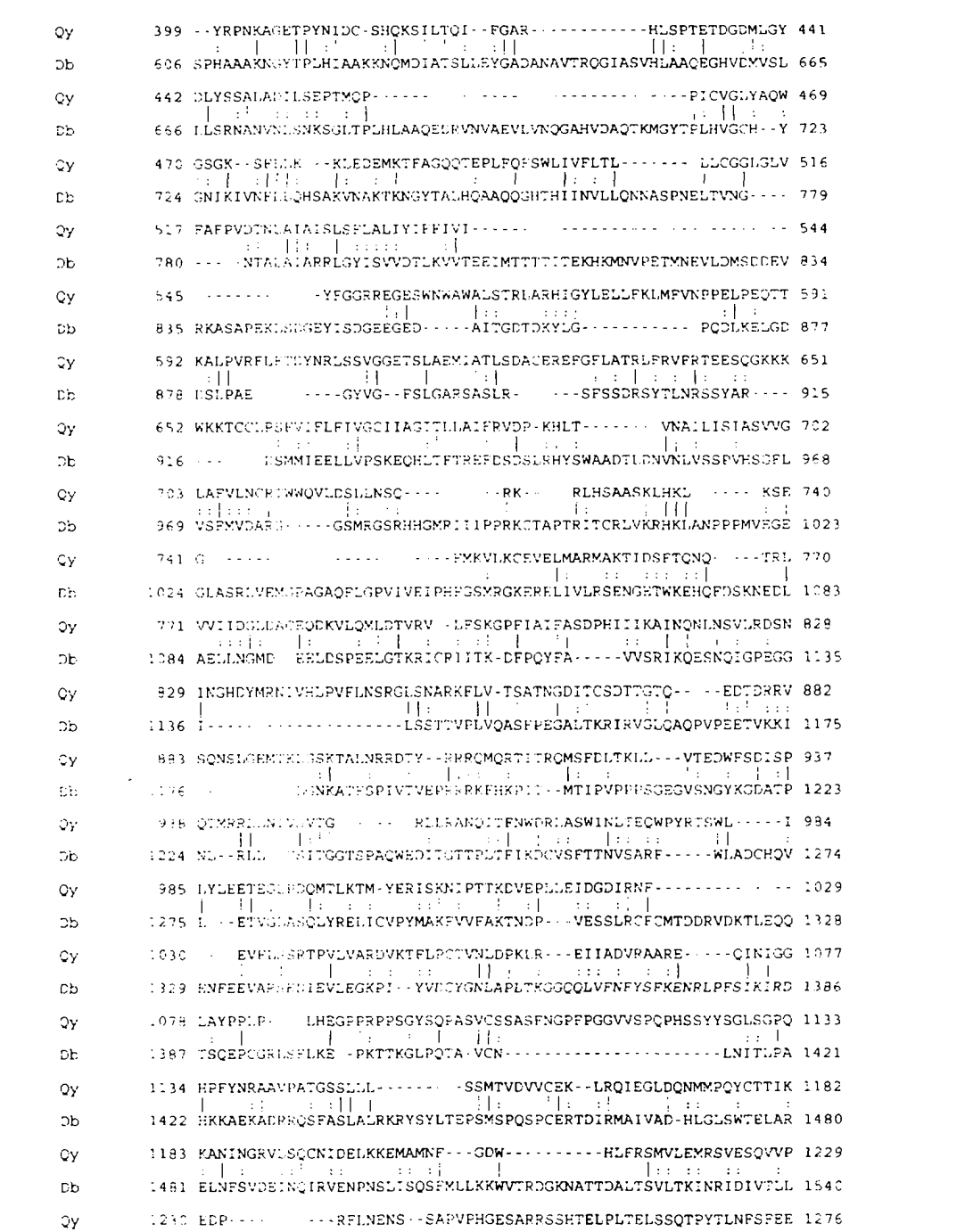


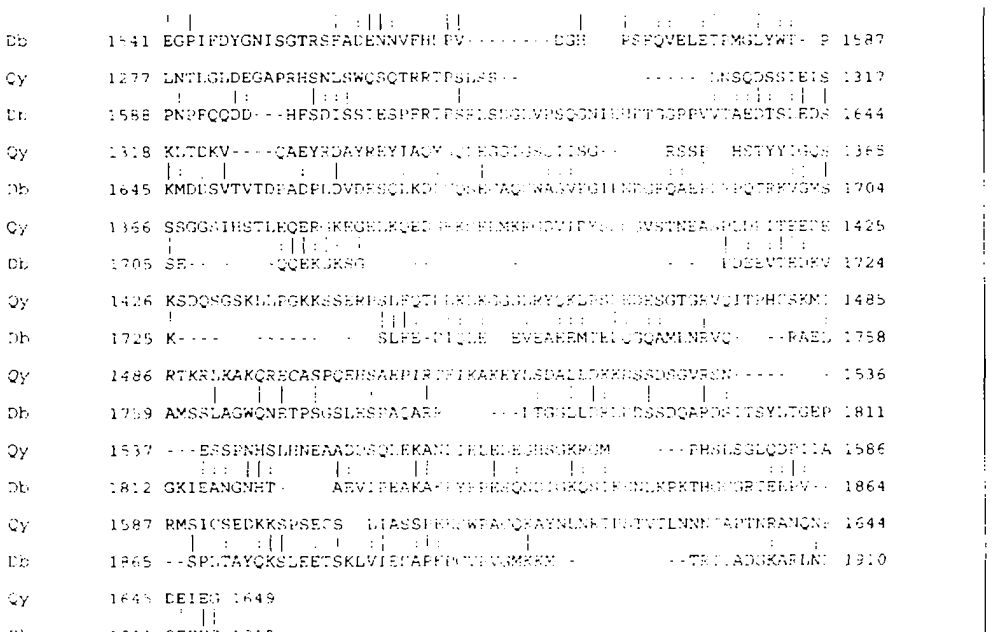




RESULT 10
T42713
ankyrin 3, splice form 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42713
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene the repeat domain.
A;Reference number: Z22237; MUID:95340433; PMID:7615634
A;Accession: T42713
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1943 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PIDN:AAB01606.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 855/1
C;Function:
A;Description: supposed to play an important role in the polarized distribution of many
A;Note: major kidney ankyrin
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 6.2%; Score 5 1; 18 2; Length 1944;
Best Local Similarity 20.2%; Pred. No. 7.2e-24;
Matches 426; Conservative 289; Mismatches 666; Indels 124; Gaps 85;

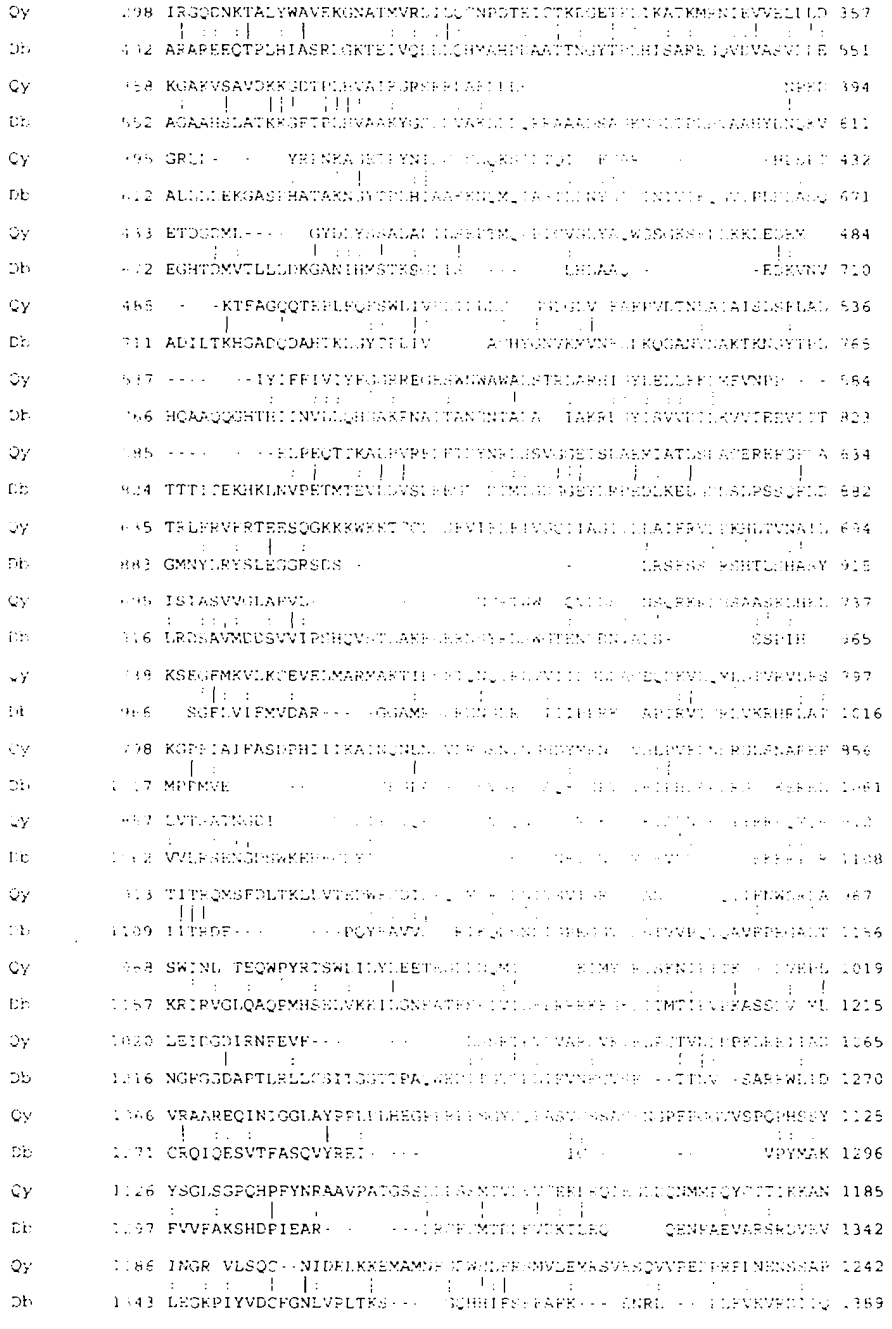


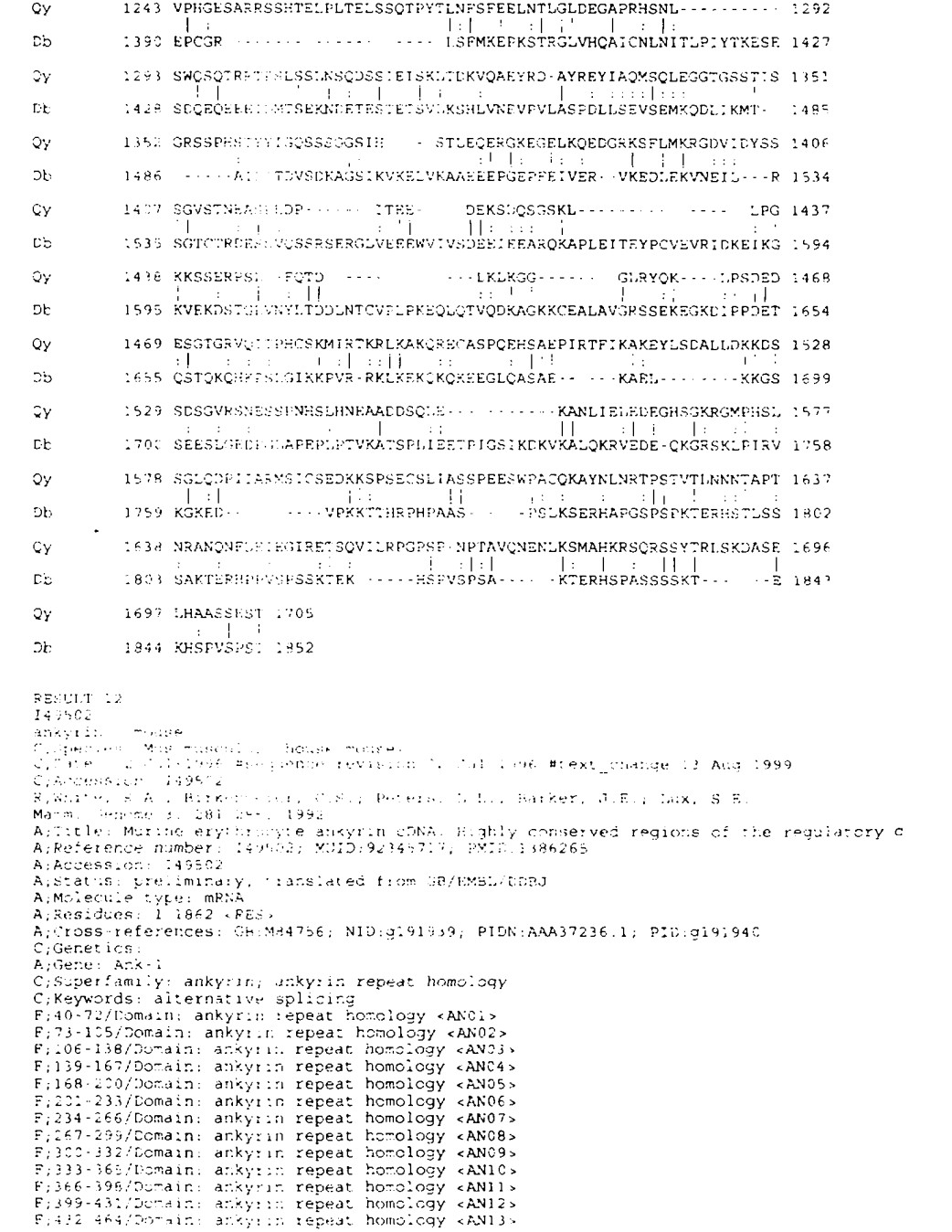


RESULT 11
S37431
ankyrin 2, neuronal long splice form - human
N;Alternate names: ankyrin B, 440K spli e form; ankyrin B, brain ankyrin, non erythroid
N;Contains: ankyrin 2, short form
C;Species: Homo sapiens (man)
C;Date: 06 Jan-1995 #sequence revision [illegible] #text change 13 Aug 1999
C;Accession: S37431; A39643; B39643; A4[illegible]; A49462; S[illegible]; S14969
R;Chan, W.
submitted to the EMBL Data Library, Sep-1994
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>
A;Cross-references: EMBL:Z26634; NID:g4[illegible]; PID:[illegible]; PID:g4[illegible]
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A;Reference number: A39643; MUID:91302476; PMID:1906[illegible]
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2077 <OTI>
A;Cross-references: GB:X56957
A;Accession: B39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443,3585-3924 <CTT>
A;Cross-references: EMBL:X56958
R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
Genomics 10, 858-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92009421; PMID:1833308
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 463-474,'PE',477-495 <TSE>
A;Cross-references: GB:M17123; NID:g178647; PIDN:AAA62828.1; PID:g178648
R;Chan, W.; Kordeli, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A;Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and
A;Reference number: A49462; MUID:94075409; PMID:8253844
A;Accession: A49462
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3924 <RES>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
C;Genetics:
A;Gene: GDB:ANK2
A;Cross-references: GDB:127607; OMIM:106410
A;Map position: 4q25-4q27
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F;2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F;63-95/Domain: ankyrin repeat homology <AN01>
F;96-128/Domain: ankyrin repeat homology <AN02>
F;129-161/Domain: ankyrin repeat homology <AN03>
F;162-190/Domain: ankyrin repeat homology <AN04>
F;191-223/Domain: ankyrin repeat homology <AN05>
F;232-264/Domain: ankyrin repeat homology <AN06>
F;265-297/Domain: ankyrin repeat homology <AN07>
F;298-330/Domain: ankyrin repeat homology <AN08>
F;331-363/Domain: ankyrin repeat homology <AN09>
F;364-396/Domain: ankyrin repeat homology <AN10>
F;397-429/Domain: ankyrin repeat homology <AN11>
F;430-462/Domain: ankyrin repeat homology <AN12>
F;463-495/Domain: ankyrin repeat homology <AN13>
F;496-528/Domain: ankyrin repeat homology <AN14>
F;529-561/Domain: ankyrin repeat homology <AN15>
F;562-594/Domain: ankyrin repeat homology <AN16>
F;595-627/Domain: ankyrin repeat homology <AN17>
F;628-660/Domain: ankyrin repeat homology <AN18>
F;661-693/Domain: ankyrin repeat homology <AN19>
F;694-726/Domain: ankyrin repeat homology <AN20>
F;727-759/Domain: ankyrin repeat homology <AN21>
F;760-792/Domain: ankyrin repeat homology <AN22>
F;795-825/Domain: ankyrin repeat homology <AN23>

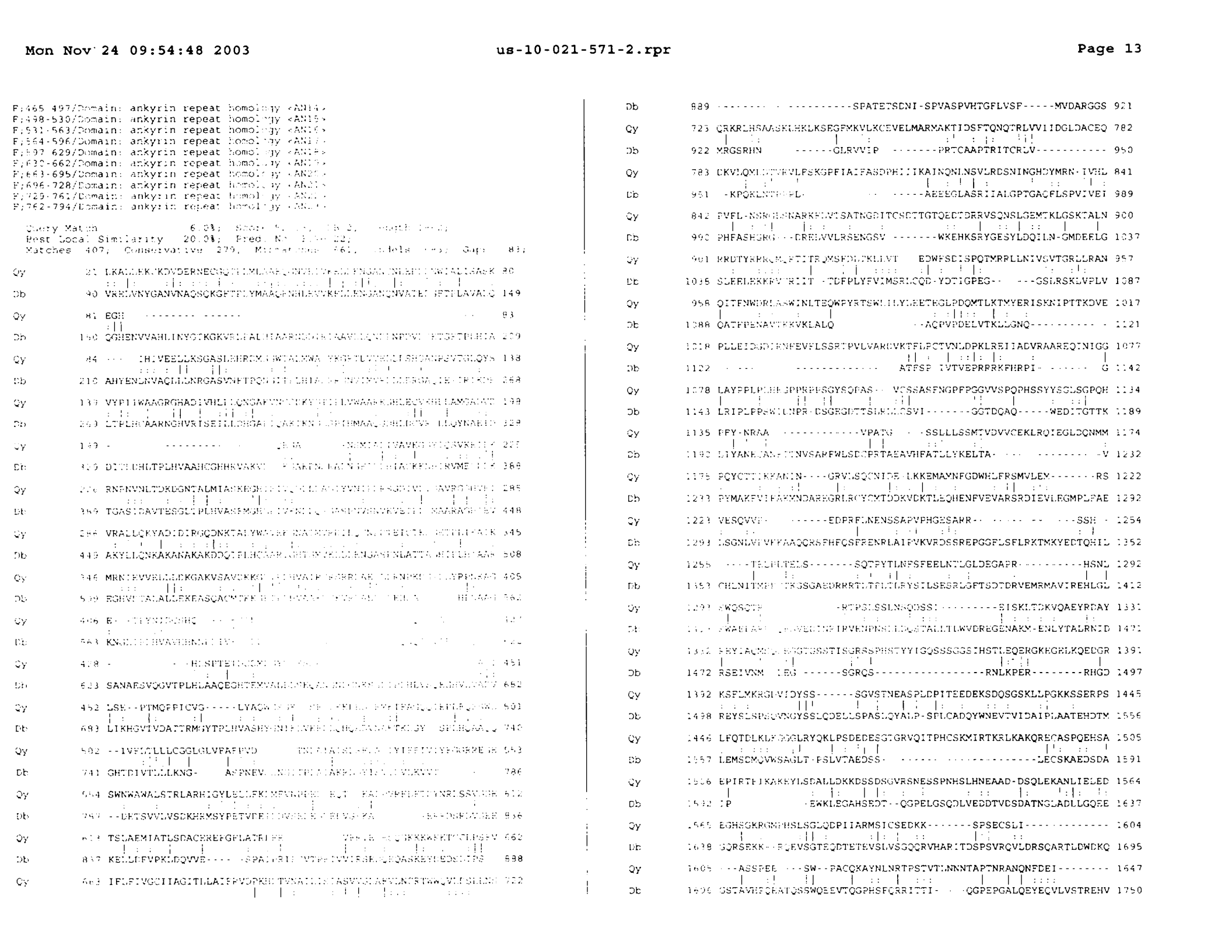
Query Match 6.2%; Score 550; DB 2; Length 3924;
Best Local Similarity 20.4%; Pred. No. 2.9e-23;
Matches 417; Conservative 279; Mismatches 727; Indels 626; Gaps 78;

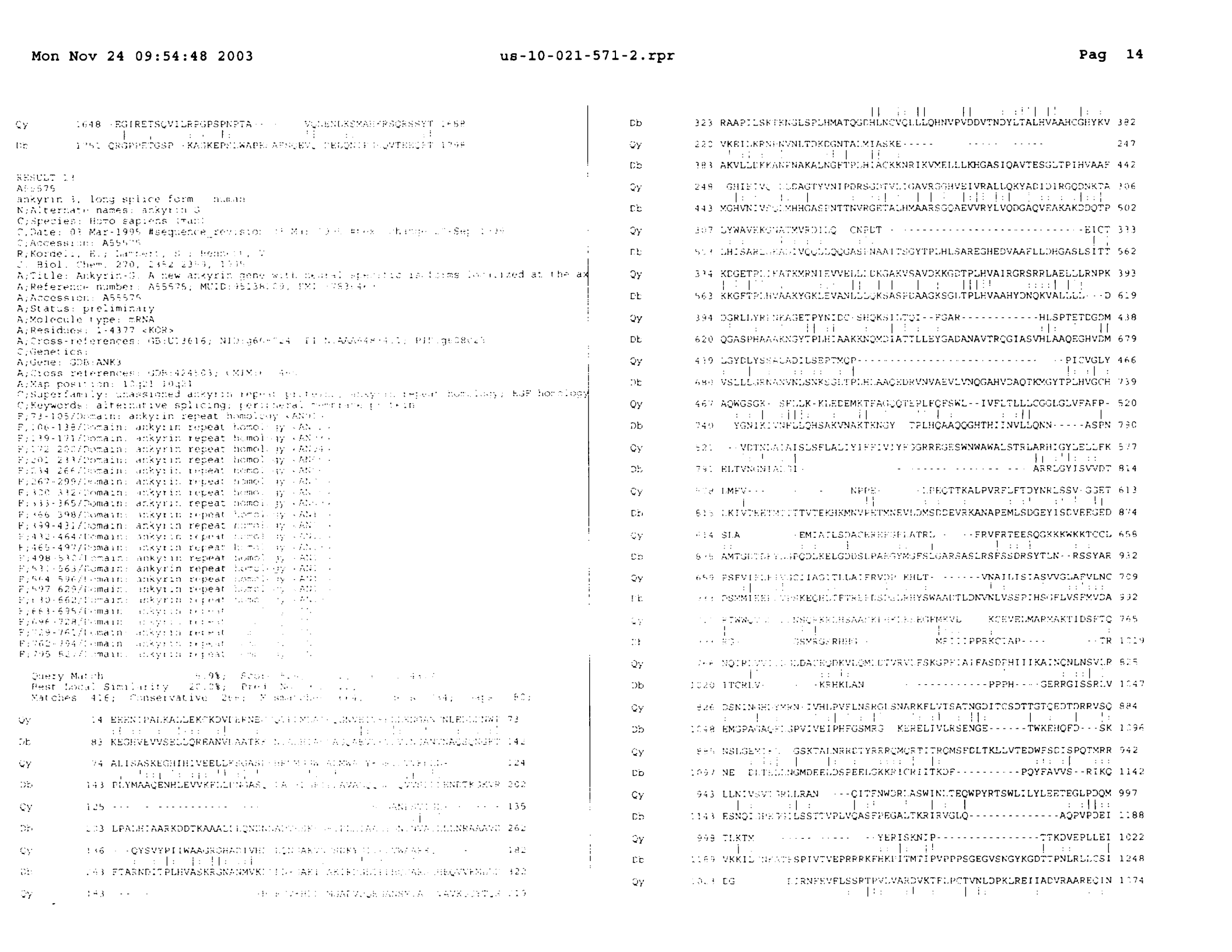


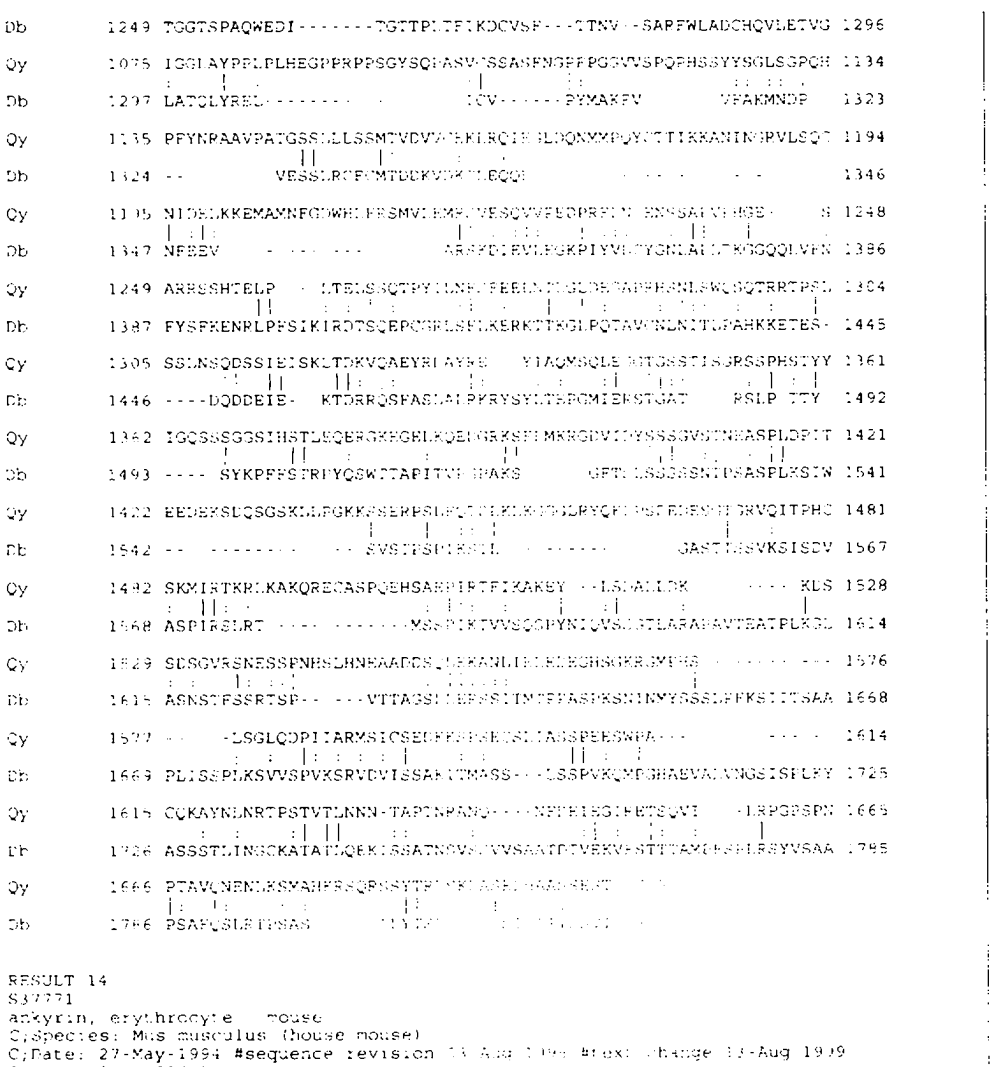




RESULT 12
I49502
Ankyrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Oct-1996 #sequence_revision 31-Jul-1998 #text_change 13-Aug-1999
C;Accession: I49502
R;Kordeli, E.; Bennett, V.; Peters, L.L., Barker, J.E.; Lux, S.E.
Mamm. Genome 5, 281-287, 1994
A;Title: Murine erythrocyte ankyrin cDNA: highly conserved regions of the regulatory d
A;Reference number: I49502; MUID:94346717; PMID:8086265
A;Accession: I49502
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1862 <PES>
A;Cross-references: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C;Genetics:
A;Gene: Ank-1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;40-72/Domain: ankyrin repeat homology <AN01>
F;73-105/Domain: ankyrin repeat homology <AN02>
F;106-138/Domain: ankyrin repeat homology <AN03>
F;139-167/Domain: ankyrin repeat homology <AN04>
F;168-200/Domain: ankyrin repeat homology <AN05>
F;201-233/Domain: ankyrin repeat homology <AN06>
F;234-266/Domain: ankyrin repeat homology <AN07>
F;267-299/Domain: ankyrin repeat homology <AN08>
F;300-332/Domain: ankyrin repeat homology <AN09>
F;333-365/Domain: ankyrin repeat homology <AN10>
F;366-398/Domain: ankyrin repeat homology <AN11>
F;399-431/Domain: ankyrin repeat homology <AN12>
F;432-464/Domain: ankyrin repeat homology <AN13>



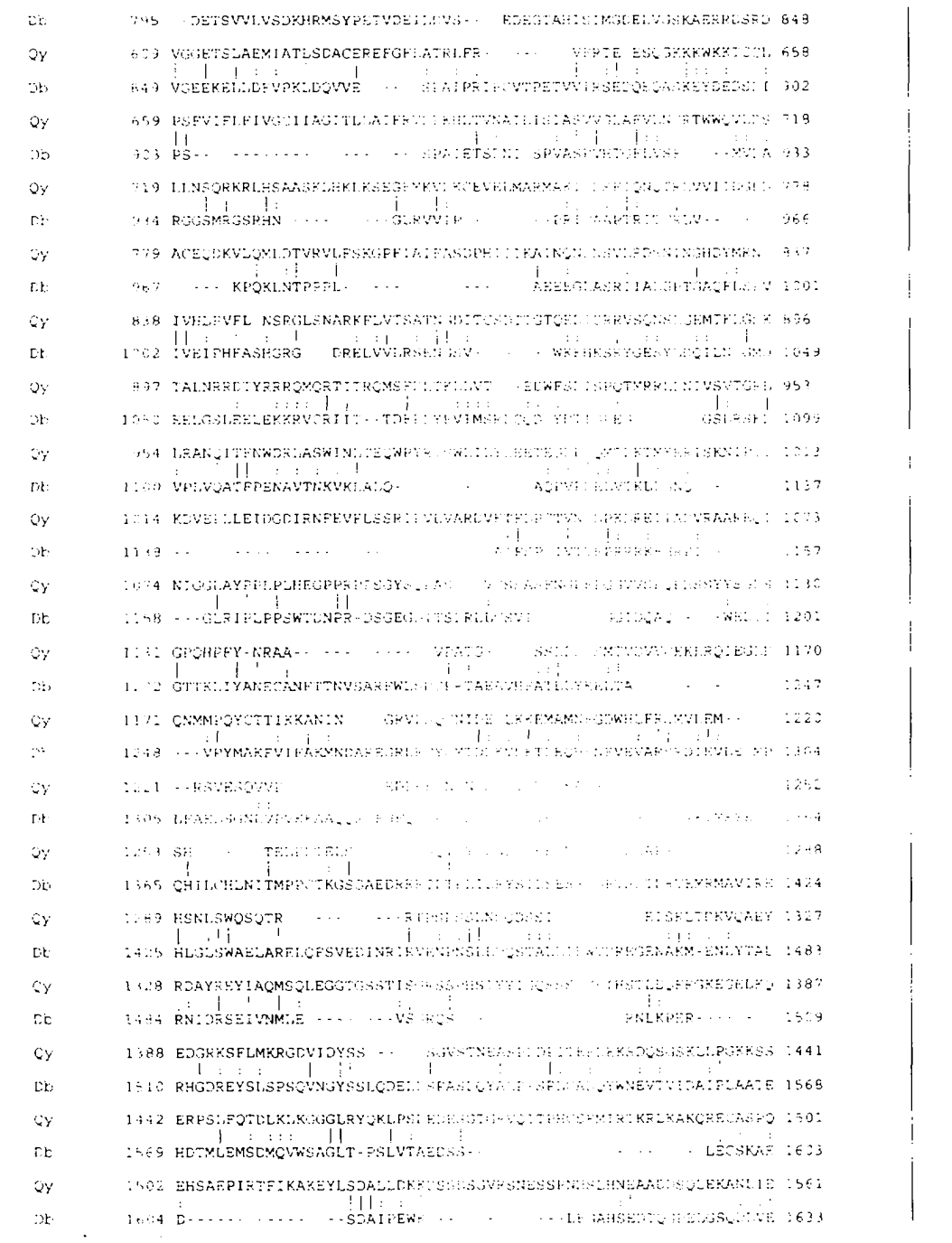




RESULT 14
S37771
ankyrin, erythrocyte - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-May-1994 #sequence_revision 19-Aug-1999 #text_change 19-Aug 1999
C;Accession: S37771
R;Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, B.J.; Cox, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found am
A;Reference number: S37771; MUID:93252845; PMID:8486647
A;Accession: S37771
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1848 <BIR>
A;Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48851.1; PID:g311817
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;48-80/Domain: ankyrin repeat homology <AN01>
F;81-113/Domain: ankyrin repeat homology <AN02>
F;114-146/Domain: ankyrin repeat homology <AN03>
F;147-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;275-307/Domain: ankyrin repeat homology <AN08>
F;308-340/Domain: ankyrin repeat homology <AN09>
F;341-373/Domain: ankyrin repeat homology <AN10>
F;374-406/Domain: ankyrin repeat homology <AN11>
F;407-439/Domain: ankyrin repeat homology <AN12>
F;440-472/Domain: ankyrin repeat homology <AN13>
F;473-505/Domain: ankyrin repeat homology <AN14>
F;506-538/Domain: ankyrin repeat homology <AN15>
F;539-571/Domain: ankyrin repeat homology <AN16>
F;572-604/Domain: ankyrin repeat homology <AN17>
F;605-637/Domain: ankyrin repeat homology <AN18>
F;638-670/Domain: ankyrin repeat homology <AN19>
F;671-703/Domain: ankyrin repeat homology <AN20>
F;704-736/Domain: ankyrin repeat homology <AN21>
F;737-769/Domain: ankyrin repeat homology <AN22>
F;770-802/Domain: ankyrin repeat homology <AN23>

Query Match 5.8%; Score 516.5; DB 2; Length 1848;
Best Local Similarity 19.6%; Pred. No. 7.8e-22;
Matches 397; Conservative 281; Mismatches 651; Indels 697; Gaps 81;

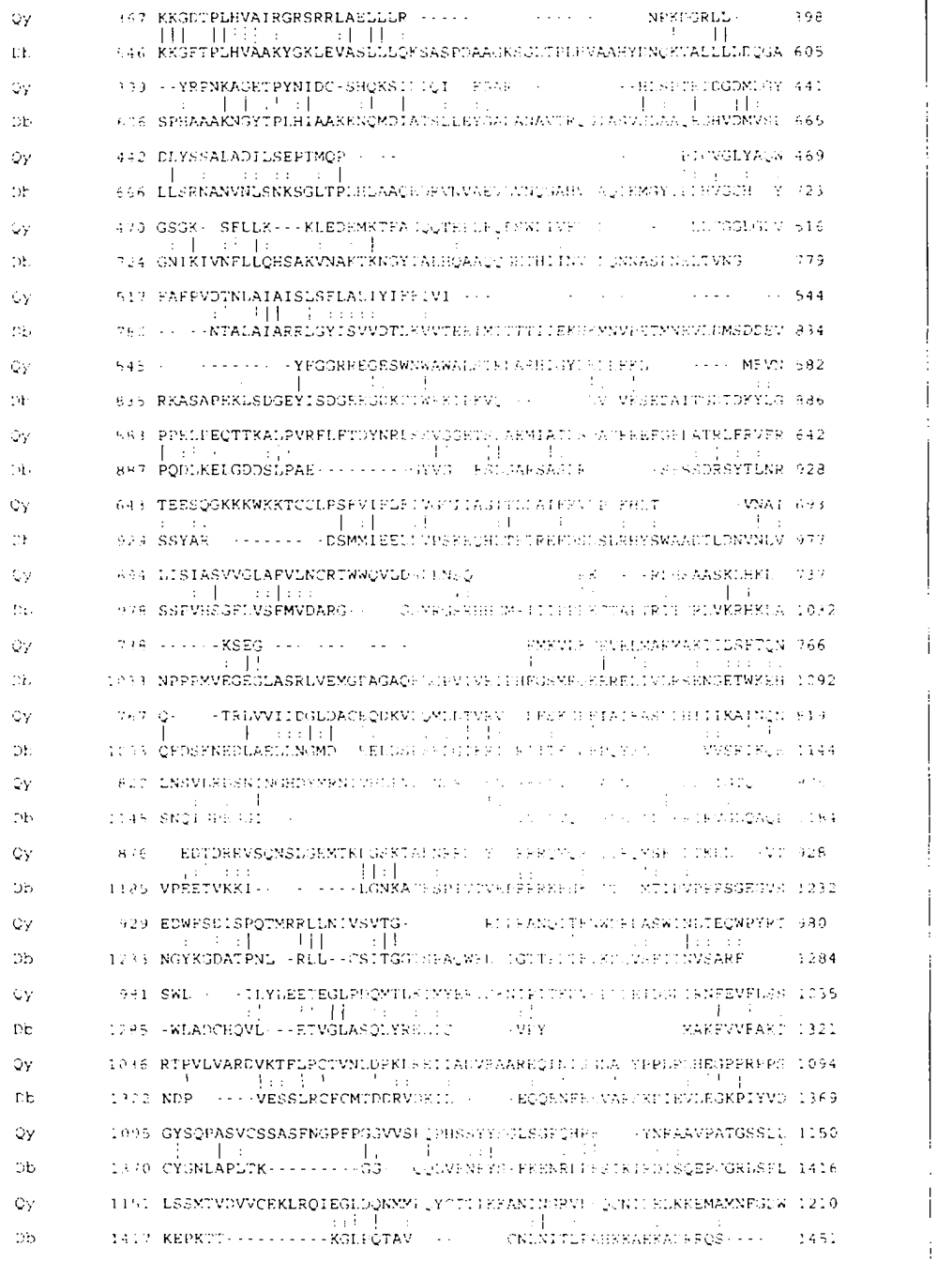


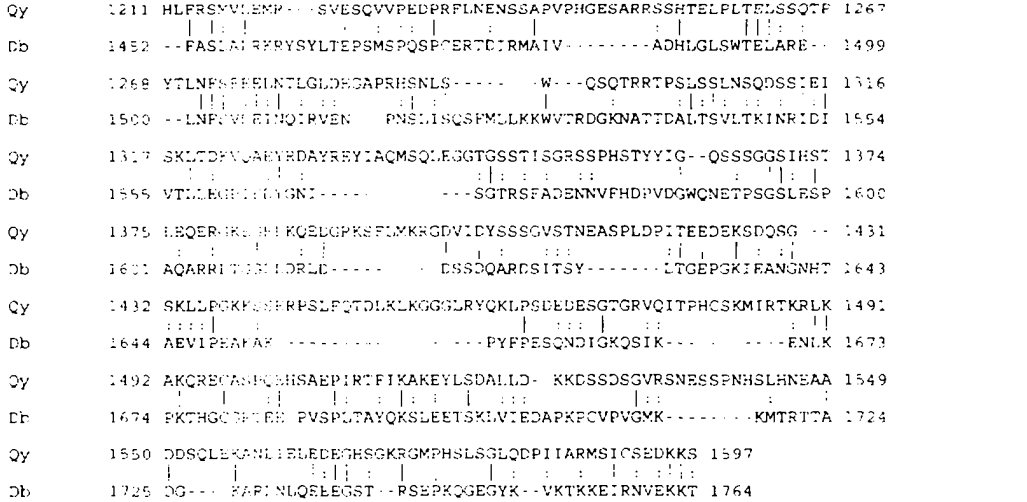




RESULT 15
T42714
ankyrin 3, splice form 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42714
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gen the repeat domain.
A;Reference number: A42187; MUID:95340633; PMID:7615634
A;Accession: T42714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <PET>
A;Cross-references: EMBL:L40633; NID:g710548; PID:g710551; PIDN:AAB01606.1
A;Experimental source: strain C57BL/6J, kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 1587/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 5.8%; Score 510.5; DB 2; Length 1765;
Best Local Similarity 20.1%; Pred. No. 1.6e-21;
Matches 396; Conservative 278; Mismatches 641; Indels 653; Gaps 78;





Search completed: November 21, 2003, 12:48:10
Job time : 40.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2003, 12:53:49 ; Search time 15.5 Seconds
(without alignments)
5201.227 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8953
Sequence: 1 MSVLISQVSNYVEENIIA ... SIRAAPV ... PRSEPVSIL 1880

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120863 seqs, 47267705 residues

Total number of hits satisfying chosen parameters: 241726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	956	6.3	1880	1	ANK1_HUMAN	P16157 homo sapien
2	950	6.2	3724	1	ANK2_HUMAN	Q12955 homo sapien
3	927.5	6.0	1862	1	ANK1_MOUSE	Q02357 mus musculu
4	526	5.9	4377	1	ANK3_HUMAN	Q12955 homo sapien
5	468	5.3	768	1	YB23_HUMAN	Q14... homo sapien
6	441.5	4.9	1059	1	Y479_HUMAN	Q15... homo sapien
7	434.5	4.9	1491	1	LARK_HUMAN	... homo sapien
8	426.5	4.8	[illegible]	1	YP1_...	... homo sapien
9	424.5	4.7	[illegible]	1	...	... homo sapien
10	420.5	4.7	1166	1	YP1_HUMAN	... homo sapien
11	418.5	4.7	824	1	ADRB_...	... mus musculu
12	360	4.1	838	1	ANK3_HUMAN	... homo sapien
13	357.5	4.0	741	1	SNSA_HUMAN	Q16... homo sapien
14	356	4.0	735	1	SNSA_MOUSE	Q9r... mus musculu
15	349.5	3.9	692	1	ANR6_HUMAN	Q9y... homo sapien
16	337	3.8	1267	1	PXT1_HUMAN	Q15... homo sapien
17	335	3.8	1401	1	DATA_RATNA	P... rattus norvegicus
18	331	3.7	525	1	ASB3_MOUSE	Q9w... mus musculu
19	329	3.7	583	1	AS15_MOUSE	Q9v... mus musculu
20	326.5	3.7	516	1	ASB3_HUMAN	Q9... homo sapien
21	328.5	3.7	747	1	V222_FOWPV	Q9... fowlpox vir
22	310.5	3.5	542	1	V155_FOWPV	Q9... fowlpox vir
23	309	3.5	668	1	V244_FOWPV	Q9... fowlpox vir
24	308	3.5	596	1	V024_FOWPV	Q9... fowlpox vir
25	307	3.5	436	1	V245_FOWPV	Q9... fowlpox vir
26	288	3.3	525	1	V228_FOWPV	Q9... fowlpox vir
27	288	3.3	1083	1	YIL2_YEAST	P... saccharomyce
28	285	3.2	1210	1	BAT8_HUMAN	Q... homo sapien
29	284	3.2	1263	1	BAT8_MOUSE	Q... mus musculu
30	271	3.1	592	1	V246_FOWPV	Q... fowlpox vir
31	270	3.0	603	1	V162_FOWPV	Q... fowlpox vir
32	265.5	3.0	434	1	AS15_HUMAN	Q... homo sapien
33	264.5	3.0	656	1	PEX1_CHIHA	P... mesocricetus
34	262.5	3.0	1150	1	ANS1_MOUSE	P59672 mus musculu
35	260	2.9	433	1	AS14_MOUSE	Q8vhs7 mus musculu
36	254.5	2.9	776	1	ANR5_HUMAN	Q9nu02 homo sapien
37	254.5	2.9	1134	1	ANS1_HUMAN	Q92625 homo sapien
38	254	2.9	318	1	ASB7_MACFA	Q9bgt9 macaca fasc
39	251	2.8	429	1	AS10_HUMAN	Q8wxi3 homo sapien
40	249.5	2.8	318	1	ASB7_MOUSE	Q91zu0 mus musculu
41	246.5	2.8	274	1	ASB7_HUMAN	Q9h672 homo sapien
42	245.5	2.8	578	1	V022_FOWPV	Q9j5h9 fowlpox vir
43	245	2.8	279	1	AS13_HUMAN	Q8wxk3 homo sapien
44	242	2.7	434	1	V023_FOWPV	Q9j5h8 fowlpox vir
45	238.5	2.7	433	1	ANR2_HUMAN	Q9gzv1 homo sapien

ALIGNMENTS

RESULT 1
ANK1_HUMAN
ID ANK1_HUMAN STANDARD; PRT; 1880 AA.
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 41, Last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE=Hematopoietic;
RX MEDLINE=90158830; PubMed=2137657;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell cycle
RT control proteins";
RL Nature 344:36-42(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175372; PubMed=1689849;
RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
RA Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN [3]
RP VARIANT HE ILE 4[illegible].
RX MEDLINE=94042[illegible]; PubMed=8[illegible];
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA Dornwell M., Herbers J., Kugler W., Ozcan R., Pekrun A.,
RA Gallagher P.G., Schroter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive
RT hereditary spherocytosis.";
RL Nat. Genet. 13:214-218(1996).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA,K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA
CC MEMBRANE.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=2.1;
CC IsoId=P16157-1; Sequence=Displayed;
CC Name=2; Synonyms=2.2;
CC IsoId=P16157-2; Sequence=VSP_000264, VSP_000265;
CC Name=3;
CC IsoId=P16157-3; Sequence=VSP_000266;

```
CC   -!- PTM: REGULATED BY PHOSPHORYLATION.
CC    . PTM: ACYLATED BY PALMITIC ACID GROUP(S).
CC   -!- DISEASE: Defects in ANK1 are a cause of hereditary spherocytosis
CC       type1 [MIM:182900]. Inheritance can be autosomal dominant or
CC       recessive.
CC    . SIMILARITY: Contains 24 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------- ... - --------  . . . . . .
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ... - --------  . . . . . .
DR   EMBL; X16609; CAA34410.1;
DR   EMBL; M28880; AAA51742.1;
DR   PIR; S28275; S3HUK.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:492; ANK1.
DR   MIM; 182900; -.
DR   GO; GO:0005200; Structural constituent of cytoskeleton; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 24.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK domain;
KW   Phosphorylation; Lipoprotein; Disease mutation; Hereditary spherocytosis;
KW   Polymorphism.
FT   INIT_MET      1      1
FT   DOMAIN        1    826       MEMBRANE DOMAIN.
FT                                -ANK N EXCHANGER PROTEIN BINDING (MAIN).
FT   DOMAIN      827   1361       CYTOPLASMIC DOMAIN (SPECTRIN BINDING) (MAIN).
FT   DOMAIN     1362   1880       REGULATORY DOMAIN.
FT                                SPECTRIN BINDING (MAIN).
FT                                NOD- AND DNA-BINDING (BY SIMILARITY).
FT   REPEAT       14     47       ANK
FT   REPEAT       51     80       ANK
FT   REPEAT      113    113       ANK
FT   REPEAT      117    146       ANK
FT   REPEAT      154    183       ANK
FT   REPEAT      187    216       ANK
FT   REPEAT      220    249       ANK 6
FT   REPEAT      303    332       ANK 7
FT   REPEAT      336    365       ANK 8
FT   REPEAT      369    398       ANK 9
FT   REPEAT      402    431       ANK 10
FT   REPEAT      435    464       ANK 11
FT   REPEAT      468    497       ANK 12
FT   REPEAT      501    530       ANK 13
FT   REPEAT      534    563       ANK 14
FT   REPEAT      567    596       ANK 15
FT   REPEAT      600    629       ANK 16
FT   REPEAT      633    662       ANK 17
FT   REPEAT      666    695       ANK 18
FT   REPEAT      699    728       ANK 19
FT   REPEAT      732    761       ANK 20
FT   REPEAT      765    794       ANK 21
FT   DOMAIN     1402   1486       DEATH.
FT   VARSPLIC   1512   1873       Missing (in isoform 2).
FT                                /FTId=VSP_000264.
FT   VARSPLIC   1874   1874       R -> S (in isoform 2).
FT                                /FTId=VSP_000265.
FT   VARSPLIC   1849   1880       TVEGPLEDPSELEVDIDYFMKHSKDHTSTPNP -> ELRGS
FT                                GLQPDLIEGRKGAQIVKRASLKRGKQ (in isoform
FT                                3).
FT                                /FTId=VSP_000266.
FT   VARIANT      20     20       R -> T.
FT                                /FTId=VAR_000595.
FT   VARIANT     462    462       V -> I (in HS).
FT                                /FTId=VAR_000596.
FT   VARIANT     618    618       R -> H (in Brueggen).
FT                                /FTId=VAR_000597.
FT   VARIANT     749    749       V -> A.
FT                                /FTId=VAR_000598.
FT   VARIANT     844    844       D -> E.
FT                                /FTId=VAR_000599.
FT   VARIANT    1285   1285       E -> D.
FT                                /FTId=VAR_000601.
FT   VARIANT    1391   1391       S -> T.
FT                                /FTId=VAR_000600.
FT   VARIANT    1691   1691       D -> N (in Duesseldorf).
FT                                /FTId=VAR_000602.
FT   VARIANT    1696   1696       R -> D.
FT                                /FTId=VAR_000603.
FT   CONFLICT    227    229       A -> S (IN REF. 2).
FT   CONFLICT   1545   1545       V -> I (IN REF. 2).
SQ   SEQUENCE   1880 AA;  206145 MW;  1C5F5E7BFD1CD428 CRC64;
```

```
  Query Match             6.4%;  Score 556;  DB 1;  Length 1880;
  Best Local Similarity   20.6%;  Pred. No. 1.2e-21;
  Matches  396;  Conservative  270;  Mismatches  635;  Indels  630;  Gaps   75;

Qy         21 LKALLEKGKDVDARNEDGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWTALISASK 80
              :: :  :: :: :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         93 VRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQ 152

Qy         81 EGH-- ---          ------------ -- - ------------ --  - --- 89
              :||
Db        153 QGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTAAVLLQNDPNPDVLSKTGFTPLHIA 212

Qy         90 -- --IHIVALLLKSGASLEHRDMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYS 138
                :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        213 AHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIE-TKTKDE 271

Qy        139 VYPIIWAAGKGHADIVHLLLQNGAKVNGSDKYGTTPLVWAARKGHLECVKHLLAMGADVD 198
             :  :   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        272 LTPLHLAAPNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGGHLDCVRLLLQYDAEID 331

Qy        199 -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
                                                                          225
Db        332 DICLEGGLTALHWAAHCGHERVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRVMELLLK 391

Qy        226 RNPNVNLETEKGNTALMIASKEGHIEVVQLLDAGTYVNIPDRSGDTVLIGAVRGGHVEI 285
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        392 TGASILAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEV 451

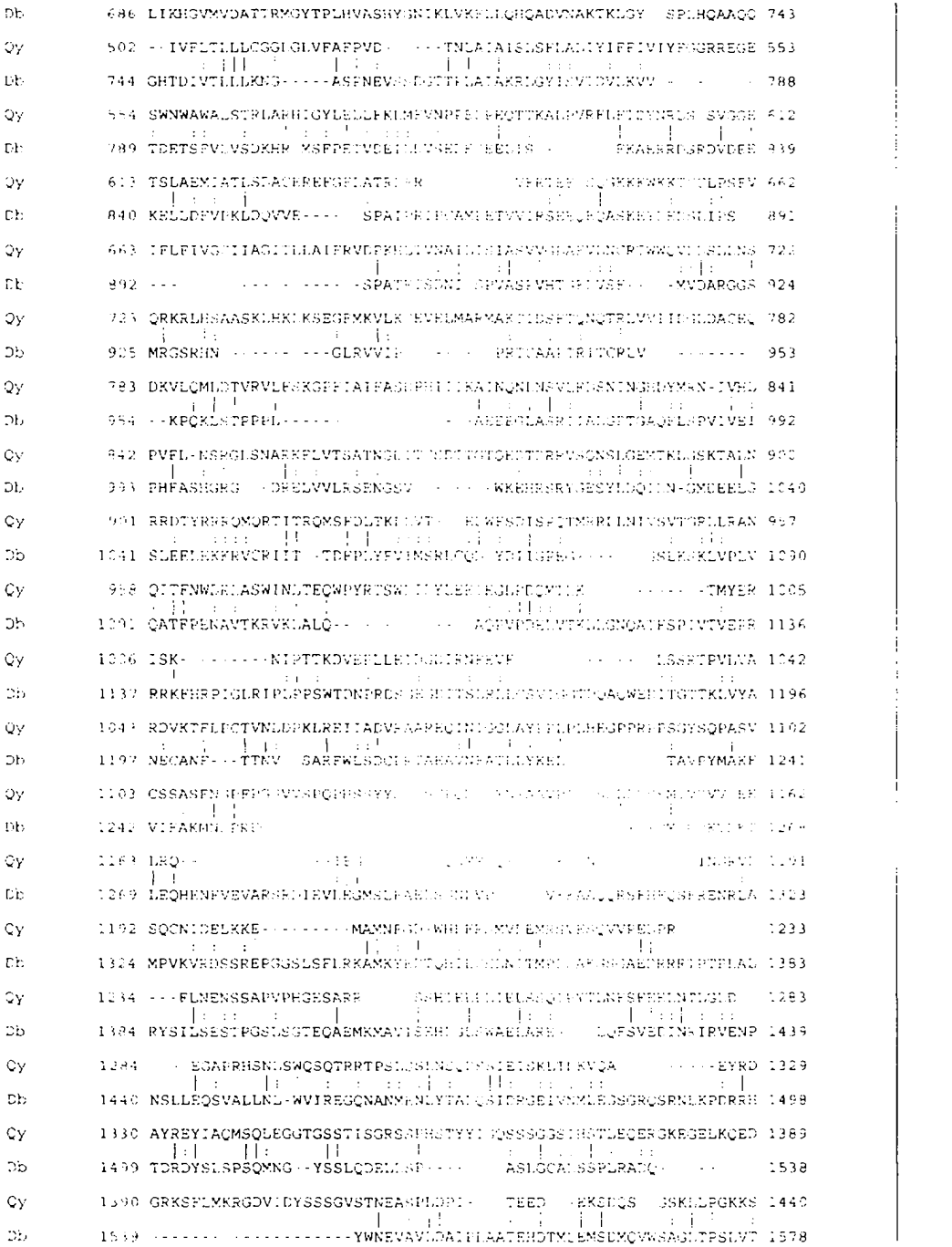
Qy        286 VRALLQKYAPDIRGQDNKTALYWAVEKGNAMVRDILQCNPDTEICTKDGETPLIKATK 345
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        452 AKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAR 511

Qy        346 MRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAG 405
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        512 EGHVEIVLALIEKEASQACMTKKGFTPLHVAAKYGKVRVAELLLE--RDA----HPNAAG 565

Qy        406 E-- TPYNIDDSHQ----KSIL-- -- -------------- -TQIFGAR------- 427
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        566 KNGLTPLHVAVHHNNLDIVKLLLPRGGSPHSPAWNGYTPLHIAAKQNQVEVARSLLQYGG 625

Qy        428 -----      --HLSPTETDGDMLGYDLYSSAMADILSEPTMQP----------PIC-- 462
                         :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        626 SANAEMQDGVTPLHLAAQEGHAEMVALLLSKQANGNLGNKSGLTPLHLVAQEGHVPVADV 685

Qy        463          -VGLYAQWGSGK--SFLLKKLED-EMKTFAGQQTEPLFQFSWL 501
                          :  :  :  :  :  :  :  :  :  :  :  :  :  :
```



Qy 1441 SERPSLFQTDLKLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIRTKRLKAKQRECASP 1500
Db 1579 AEDSSL- -------- -ECSKAEDSDATGH-- --------EWKLEGALSEEPRGP 1612

Qy 1501 QEHSAEPKTPIKAKEYLSDA---LLDKKDSSDSGVRSNESSPNHSLHNEAADDSQLEKA 1557
Db 1613 ELGSLE- -VEDDTVDSDATNGLIDLLF-QEEGQRSEEKLPGSKRQDDATGAGQ-DSE 1666

Qy 1558 NLIELEDHGN-KRGMPHSLSGLQCPIIAPMSICSEDK ----KSPSECSLIASSPEESW 1612
Db 1667 NEVSLVSGD-VSQARITHS --PTVSQVTERSQDRLQDWDADGSIVSYLQDAAQGSW 1720

RESULT 2
ANK2_HUMAN
ID ANK2_HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RT ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN [2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY 1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075403; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
RT domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [4]
RP SEQUENCE OF 4[illegible]4 FROM N.A.
RX MEDLINE=97006921; PubMed=8939908;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
RT ankyrin gene.";
RL Genomics 10:858-866(1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q01484-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q01484-2; Sequence=VSP_000267, VSP_000268;
CC Name=3;
CC IsoId=Q01484-3; Sequence=VSP_000268;
CC -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC AND FUNCTION (POTENTIAL).
CC -!- SIMILARITY: Contains 23 ANK repeats.
CC -!- SIMILARITY: Contains 1 death domain.
CC ---------- ---- ----------

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -!- ------ ----- - -------- ... . .. -----
DR   EMBL; X66957; CAA47278.1; -.
DR   EMBL; X86368; CAA60277.1; -.
DR   EMBL; Z25634; CAA82644.1; -.
DR   EMBL; M97123; AAA62826.1; -.
DR   PIR; S33431; S33431.
DR   HSSP; P42771; 1D2Z.
DR   Genew; HGNC:6403; ANK2.
DR   MIM; 106410; -.
DR   InterPro; IPR002110; Ank.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000936; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Phosphorylation.
FT   REPEAT       67     97       ANK .
FT   REPEAT       98     128      ANK .
FT   REPEAT      129     159      ANK .
FT   REPEAT      161     191      ANK .
FT   REPEAT      193     223      ANK .
FT   REPEAT      242     272      ANK .
FT   REPEAT      275     304      ANK .
FT   REPEAT      308     337      ANK .
FT   REPEAT      341     370      ANK .
FT   REPEAT      374     403      ANK .
FT   REPEAT      407     436      ANK .
FT   REPEAT      440     469      ANK .
FT   REPEAT      473     502      ANK .
FT   REPEAT      506     535      ANK .
FT   REPEAT      539     568      ANK .
FT   REPEAT      572     601      ANK .
FT   REPEAT      605     634      ANK .
FT   REPEAT      638     667      ANK .
FT   REPEAT      671     700      ANK .
FT   REPEAT      704     733      ANK .
FT   REPEAT      737     766      ANK .
FT   REPEAT      770     799      ANK .
FT   REPEAT      803     832      ANK .
FT   DOMAIN     1773    1950      REPEAT REGION.
FT   REPEAT     1775    1784      REPEAT A.
FT   REPEAT     1785    1796      REPEAT A.
FT   REPEAT     1797    1808      REPEAT A.
FT   REPEAT     1819    1829      REPEAT A.
FT   REPEAT     1821    1832      REPEAT A.
FT   REPEAT     1833    1844      REPEAT A.
FT   REPEAT     1845    1856      REPEAT A.
FT   REPEAT     1857    1867      REPEAT A (APPROXIMATE).
FT   REPEAT     1868    1879      REPEAT A.
FT   REPEAT     1880    1891      REPEAT A.
FT   REPEAT     1892    1902      REPEAT A (APPROXIMATE).
FT   REPEAT     1903    1914      REPEAT A.
FT   REPEAT     1915    1926      REPEAT A.
FT   REPEAT     1927    1938      REPEAT A.
FT   REPEAT     1939    1950      REPEAT A.
FT   DOMAIN     3536    3620      DEATH
FT   VARSPLIC   1039    1039      Q -> QFLGKIHLPIAIPPDLNEGSSLVSRITQLGPIGTP
FT                                (in isoform 2).
FT                                /FTId=VSP_000267.
FT   VARSPLIC   1444    1628      Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_000268.
FT   CONFLICT    475     476      GQ -> PE (IN REF. 4).
FT   CONFLICT    971     971      I -> S (IN REF. 1).
FT   CONFLICT   3581    3582      QY -> HA (IN REF. 1).
FT   CONFLICT   3586    3586      I -> Y (IN REF. 1).
SQ   SEQUENCE   3924 AA;  440337 MW;  52AC496C428E29D2 CRC64;

  Query Match             6.2%;  Score 550;  DB 1;  Length 3924;
  Best Local Similarity   30.4%;  Pred. No. 8.2e-21;
  Matches  417;  Conservative  270;  Mismatches  727;  Indels  626;  Gaps   78;

Qy         14 EEENIPALKAYLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWT  73
              :| :: :  :  :  ||   : | . | : :  |  :||  :  |||   :  : :|
Db         73 KEGHVGLVQRLLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFT 132

Qy         74 ALISASKEGHINIVEELLKSGAS---------- ---------------LEHRDMG -- 103
              |  |::|    :|: ||:: |:                              :|:   '
Db        133 PLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVR 192

Qy        104 ------      -  ----------------  -----GWTALMWACYKGRTDVVELL 123
                                                    '|:|  :  | :  :|  |:
Db        193 LPALHIAARKNDTKSAALLLQNDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLL 252

Qy        124 LSHGANPNVTDLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKG- 182
              : ||    :  :  :  :  :|: :: |:  ! :::   : ; ||  !| | |
Db        253 LNRGAAVTPTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 311

Qy        183 ------      -  ------HLECVKHLLAMGADVDQEGANSMTALIV 210
                             |:;||||;   | :!    : :|| |
Db        312 DQVVELLIERGAPLLARTKNGLSPLHMAAQGDHVECVKHLLQHKAPVDDVTLDYLTALHV 371

Qy        211 AVKGGYTDSVKELLKRNPNVNLTDKDGNTALMIASKE- ---------------------- 247
              :   |:    :  :| :  | |   :|  | |!  :
Db        372 AAHCGHYEVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGASIQAITESG 431

Qy        248  --------  HIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADID 297
                       | : ||   |      :: :  |:    | :  ||: | ||:    :|
Db        432 LTPIHVAAFMGHLNIVILLLQNGASPDVTNIRGETALHMAARAGQVEVVRCLLRNGALVD 491

Qy        298 IRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLD 357
              |  :  :  |:  '  |   :|: :||     :  ' :| ||  :  :    :|  :  :
Db        492 ARAREEQTPLHIASRLGKTEIVQLLLQHNAHPDAATTNGYTPLHISAREGQVDVASVLLE 551

Qy        358 KGAKVSAVDKFGDTPLHVAIRGRSRPLAELLLR----------   - ---------NPKD 394
              |  :   ||  |||  :    :|: ' :                             |  :
Db        552 AGAAHSLLARPSETPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHVAAHYDNQKV 611

Qy        395 GRLL       YRENKAQETPYNIDT SEQKSILTQI--FGAR  ----       HLSPT 432
                 ,'      :  | |  :  :|  :  :  :  :||                 ||:
Db        612 ALLLLEKASQHATAKNGYTPLHIAAKKNQMQIASTLLNYGAEINIVTKQGVTPLHLASQ 671

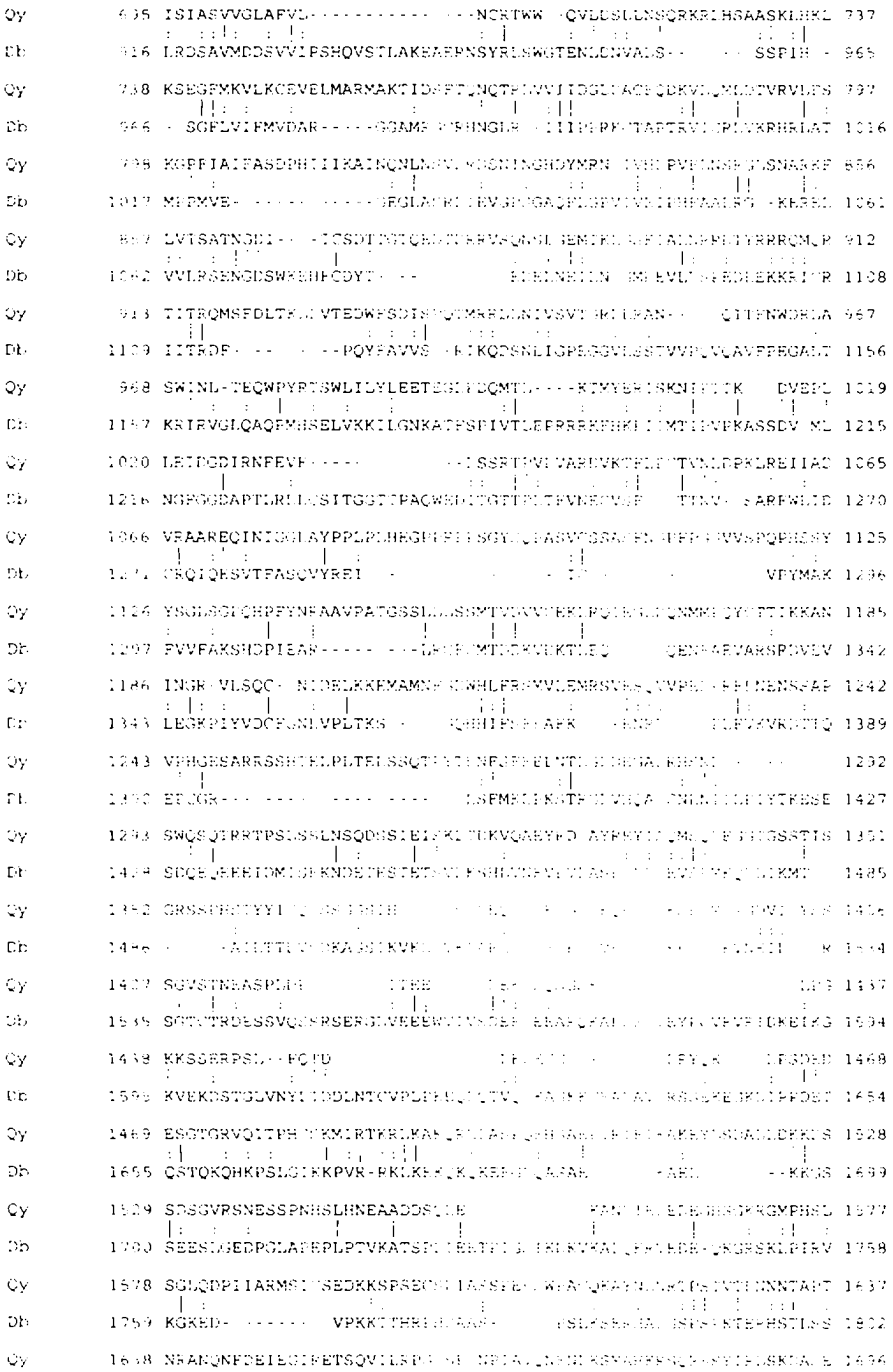
Qy        433 ETDGDML  ----GYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKLEDEM-- 484
              |    |:        ::: |  : : |          : | ||             ||::
Db        672 EGHTDMVTLLLDKGANIHMSTKSGLTS- -------LHLAAQ------------EDKVNV 710

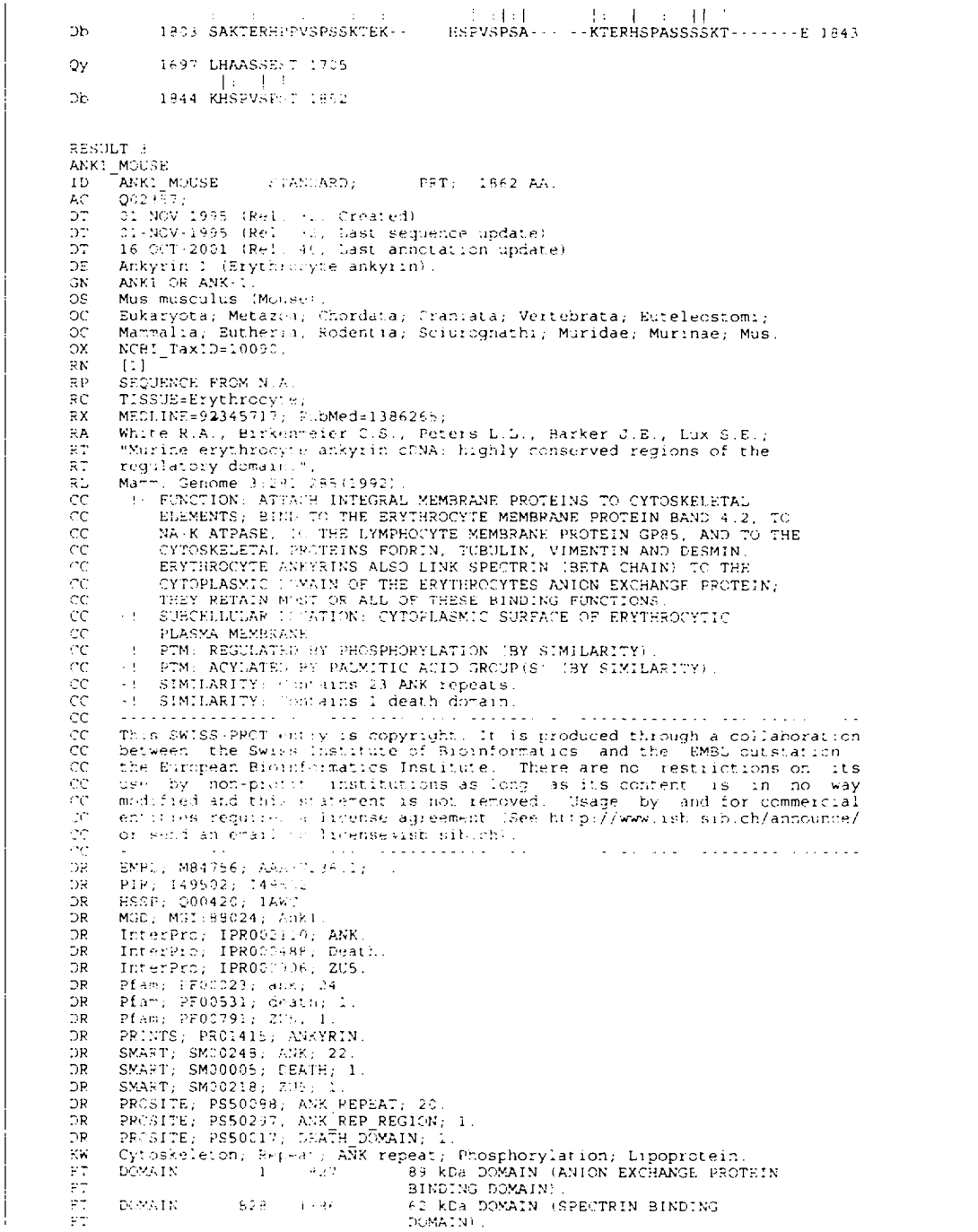
Qy        485 ---  KTPAMGQTEPLFQFSWLIVFLTLLLC--GGLGLV-FAFPVDTNLAIAISLSPLAL 536
                   |                :: |||      ;  | : :| |       |:       : ,
Db        711 ADILTKGAEQDAHTKLGYTPLIV-----ACHYGNVKMVNFLLKQGANVNAKTKNGYTPL 765

Qy        537 -----   IYIPPIVIYFGGRREGESWNKAWALSTRLARHIGYLELLFKLMPVNPP---- 584
                      :  :::  :  ,  :  ||:  :|: :||: :: |  :
Db        766 HQAAQQGHTHIINVLLQHGAKPNATTANGNTALA--IAKRLGYISVVDTLKVVTEEVTTT 823

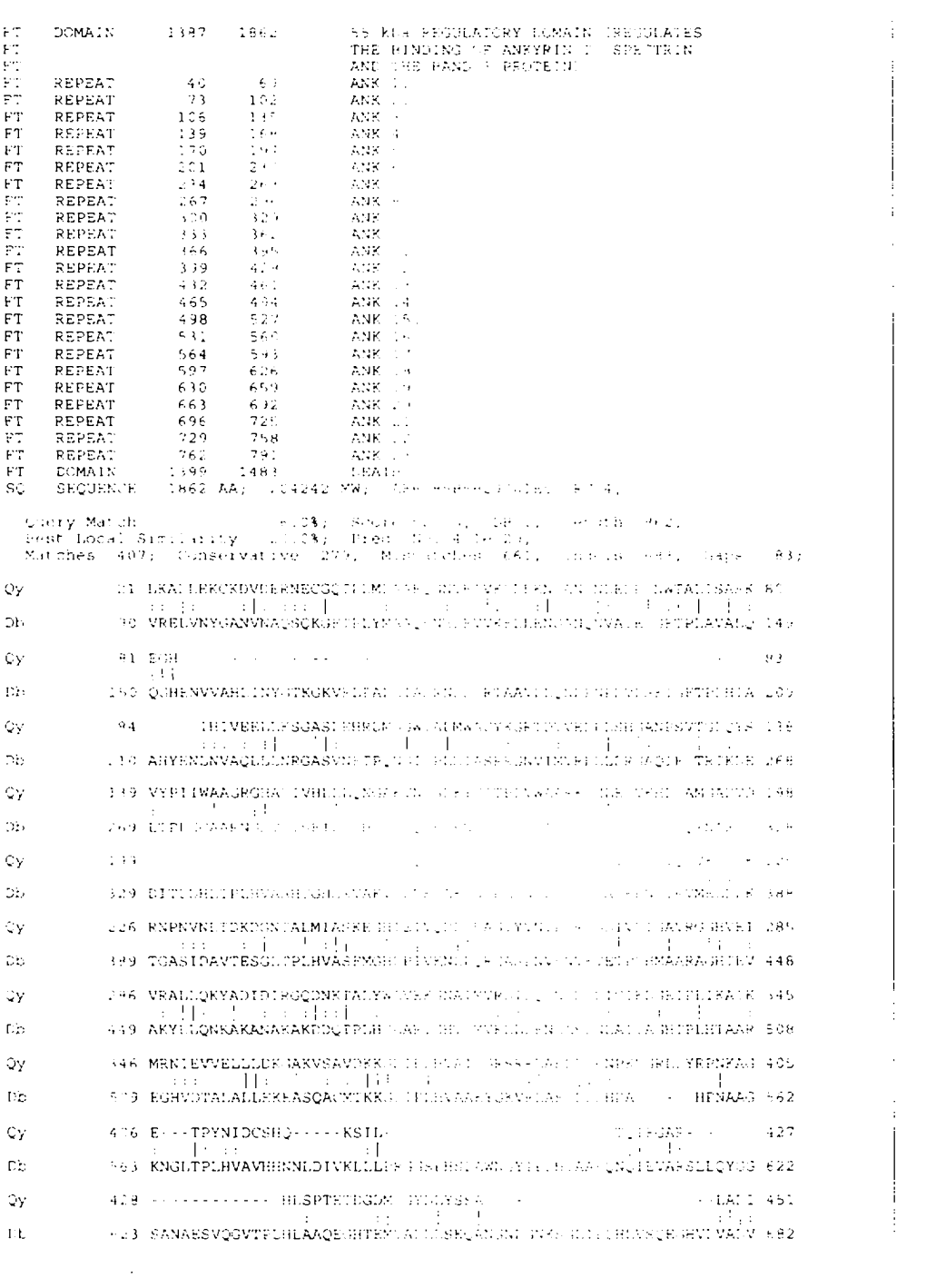
Qy        585 ------  --ELPEQTTKALPVRFLFTDYNRLSSVGGETSLAEMIATLSDACEREFGFLA 634
                       :|  |: : |     | : ::  :|:    ;  :  ' |       ||
Db        824 TTTITEKDKLNVPETMTEVLDVSDEEGD-DTMTGDGGEYLRPEDLKELGDDSLPSSQFLD 882

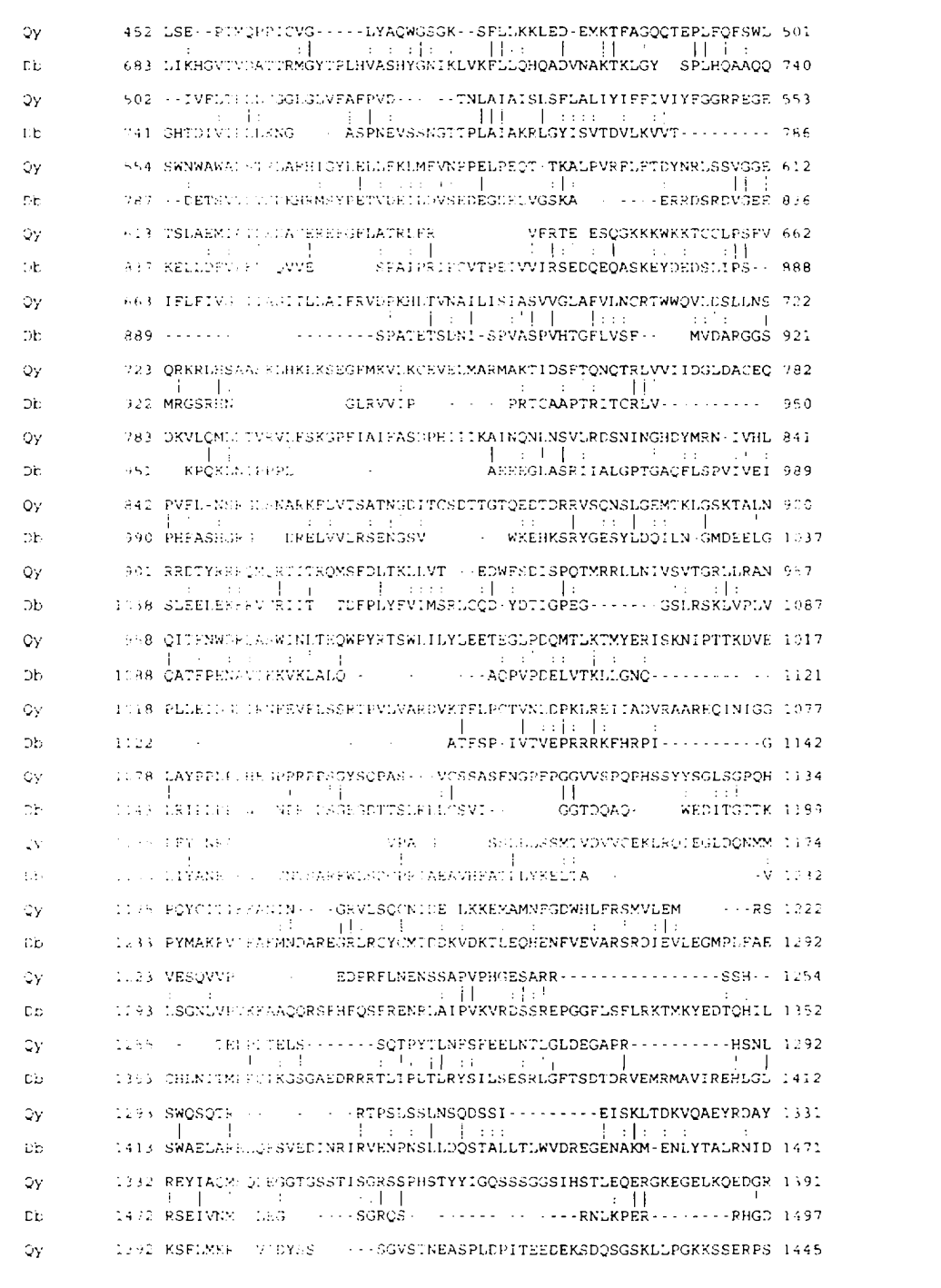
Qy        635 TRLFRVFDTDRSQGKKKWKKTCCLPSFVIFLFIVGCIIAGITLLAIPRVDPKHLTVNAIL 694
                  :  :                                          :  |  !  |  :
Db        883 GMNYLRYSLEKGRSDS----  -      ------ ---------LRSFSSDRSHTLSHASY 915
```

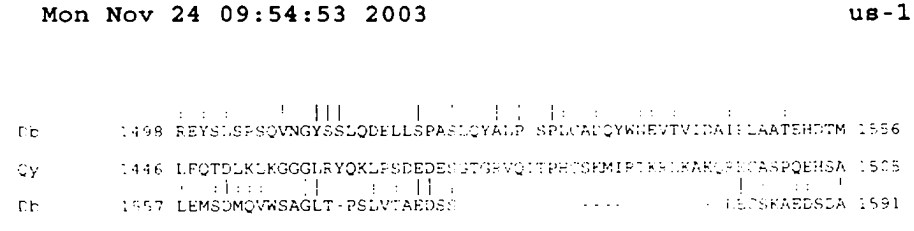




```
RESULT 3
ANK1_MOUSE
ID   ANK1_MOUSE     STANDARD;      PRT;  1862 AA.
AC   Q02357;
DT   01-NOV-1995 (Rel. 31, Created)
DT   01-NOV-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ankyrin 1 (Erythrocyte ankyrin).
GN   ANK1 OR ANK-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=92345717; PubMed=1386265;
RA   White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT   "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT   regulatory domain.",
RL   Mamm. Genome 3:241-285(1992).
CC   -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC       ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC       NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC       CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC       ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC       CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC       THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC       PLASMA MEMBRANE.
CC   -!- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).
CC   -!- PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M84756; AAA37616.1; -.
DR   PIR; I49502; I49502.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:89024; Ank1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
FT   DOMAIN        1    827       89 kDa DOMAIN (ANION EXCHANGE PROTEIN
FT                                BINDING DOMAIN).
FT   DOMAIN      828   1-##       62 kDa DOMAIN (SPECTRIN BINDING
FT                                DOMAIN).
```







```
Db       1498 REYSLSPSQVNGYSSLQDELLSPASLQYALP SPLVALQYWREVTVIDAIELAATEHDTM 1556

Qy       1446 LFQTDLKLKGGGLRYQKLPSDEDENSTGRVQITPHTSFMIPTKSLKAKLPICASPQEHSA 1505

Db       1557 LEMSDMQVWSAGLT-PSLVTAEDSS             ----          - IETSKAEDSSA 1591

Qy       1506 SPIRTFIKAKEYLSDALDLKKDSSHIVFSNE-SPMRSLHD-AAT PSIIEFANLIELHT 1564

Db       1592 IP------ -- - EWKLESNAISE  QSRET INQKQVE  VCVLSDAIVHLADLLSQEF 1637

Qy       1565 EGHSGKRGMPHSLSGLQDFIARMKISEGKK -      SPLSTSLI-      -      - 1604

Db       1638 GQRSEKK  RQEVSGTEQDIFTEVLVSLQRVHAFIISFSVRLVLEFQARILMWTKQ 1695

Qy       1605 ---ASSPEE-- -SW  PAVQKAYLNRTPSTVTINNITAITNRANQNSTST  ----- 1647

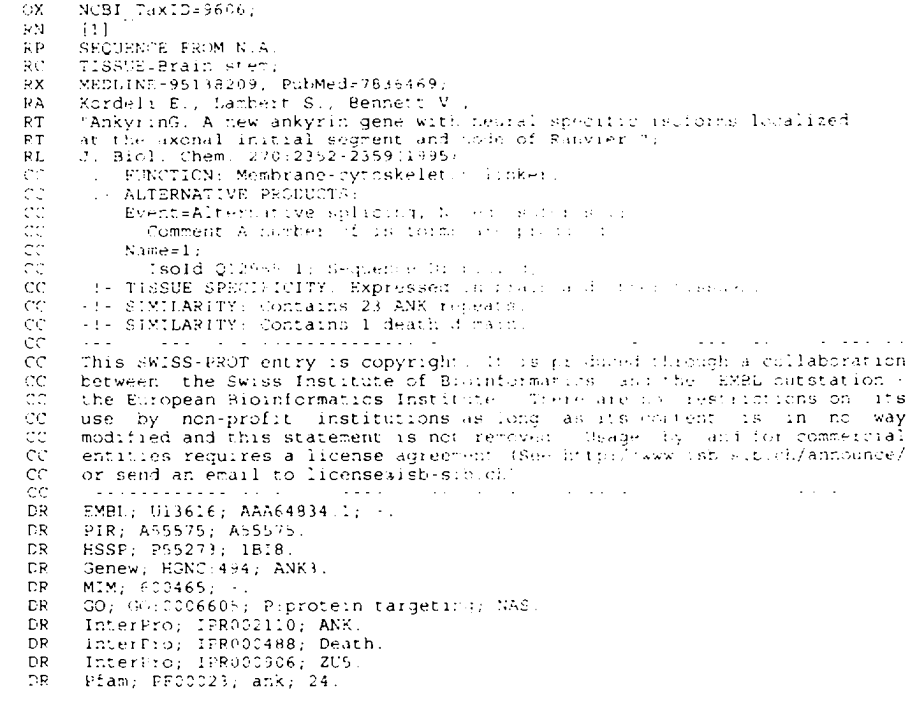
Db       1696 GSTAVHPQEATQSSWQEEVTQGPHIPQRRITTI     QSIEPIALQEYEQVIVSTREHV 1750

Qy       1648 -EGIRETSQVIIRPGPSPNPTA--  --    VQNENIKSMAHKSSQPSSYT 1698

Db       1751 QRGPPETDSP--KAGKEPSLWAPESAFSQEVQGGETQNIPGEQVTEEQFT 1798
```

RESULT 4
ANK3_HUMAN
ID ANK3_HUMAN STANDARD; PRT; 4377 AA.
AC Q12955;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem;
RX MEDLINE=95148209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.,
RT "AnkyrinG. A new ankyrin gene with neural specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995);
CC -!- FUNCTION: Membrane-cytoskeleton linker.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; [illegible]
CC Comment=A number of isoforms are produced.
CC Name=1;
CC Isold=Q12955-1; Sequence=Displayed;
CC -!- TISSUE SPECIFICITY: Expressed in brain and [illegible].
CC -!- SIMILARITY: Contains 23 ANK repeats.
CC -!- SIMILARITY: Contains 1 death domain.
CC ---------- [illegible]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; U13616; AAA64834.1; -.
DR PIR; A55575; A55575.
DR HSSP; P55273; 1BI8.
DR Genew; HGNC:494; ANK3.
DR MIM; 600465; -.
DR GO; GO:0006605; P:protein targeting; NAS.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102 ANK 1.
FT REPEAT 106 135 ANK 2.
FT REPEAT 139 168 ANK 3.
FT REPEAT 172 201 ANK 4.
FT REPEAT 203 230 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 267 296 ANK 7.
FT REPEAT 300 329 ANK 8.
FT REPEAT 333 362 ANK 9.
FT REPEAT 366 395 ANK 10.
FT REPEAT 399 428 ANK 11.
FT REPEAT 432 461 ANK 12.
FT REPEAT 465 494 ANK 13.
FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.
FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT DOMAIN 1519 1598 SER-RICH.
FT DOMAIN 4098 4174 DEATH.
SQ SEQUENCE 4377 AA; 480389 MW; F42379E55768B684 CRC64;

Query Match 5.9%; Score 526; DB 1; Length 4377;
Best Local Similarity 20.0%; Pred. No. 1.9e-19;
Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;



```
Qy         14 EEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWT 73

Db         83 KEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFT 142

Qy         74 ALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVVELLL------- 124

Db        143 PLYMAACEGHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVR 202

Qy        125 ------              ---------- - -- -SHGANPSVTGL ------- 135

Db        203 LPALHIAA-RYDNKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVD 262

Qy        136 --QYSVYHIAAAGRGHADIVRLLLQNGAKVNCSDKYGTTPLVWAARKG--------- 182

Db        263 FTARNDITGLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLD 322

Qy        183 ------ - -  - ---------HLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQS 219

Db        323 RAAPILSKTNGLSPLHMATQGDHLNCVQLLLQHNVPVDDVTNDYLTALHVAAHCGHYKV 382

Qy        220 VKEILKRNPNVNLTDKDGNTALMIASKE- ---------------------------- 247

Db        383 AKVLLDKKANPNAKALNGFTPLHIACKKNRIKVMELLLKHGASIQAVTESGLTPIHVAAF 442

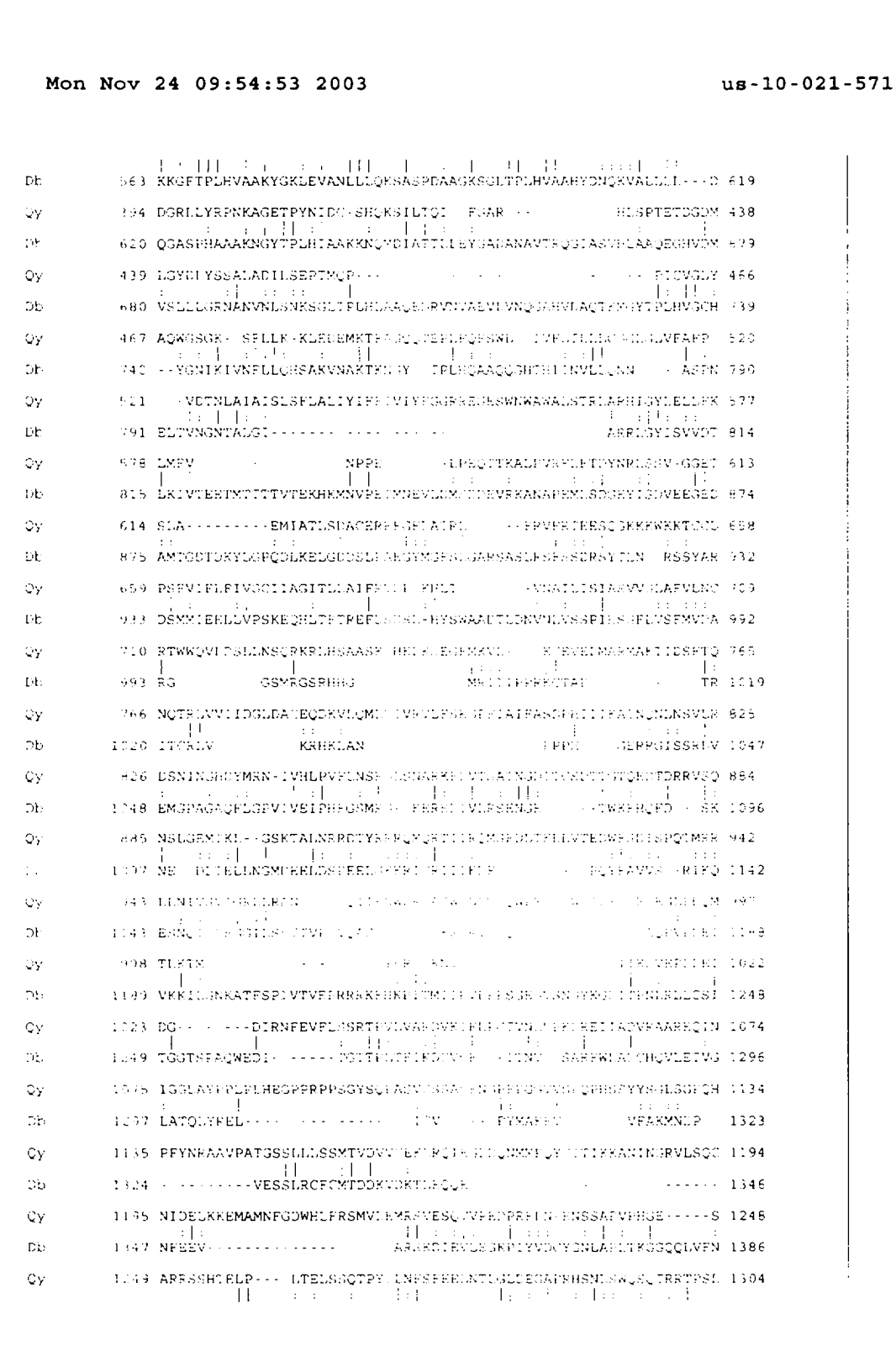
Qy        248  GHIEIVQLLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQDNKTA 306

Db        443 MGHVNIVSQLMHHGASPNTTNVRGETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTP 502

Qy        307 LYWAVEKGNATMVRDILQ- CNPDT---- ------------------------EICT 333

Db        503 LHISARLGRADIVQQLLQQGASPNAATTSGYTPLHLSAREGHEDVAAFLLDHGASLSITT 562

Qy        334 KDGETPLDIAEKMRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRRLAELLLRNPK 393
```



```
Db        563 KKGFTPLHVAAKYGKLEVANLLLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLL---D 619

Qy        394 DGRLLYRPNKAGETPYNIDV-SHQKSILTQI  FNAR --          HLSPTETDGDM 438

Db        620 QGASPHAAAKNGYTPLHIAAKKNQVDIATTLLEYGADANAVTFQGISAVHLAAQEGHVDM 679

Qy        439 LGYDIYSSADADILSEPTMQP---                -    -- PICVGDY 466

Db        680 VSLLLGRNANVNLSNKSGLTPLHLAAQEGHVDIAEVLVNQGAHVLAQTKNGYTPLHVGCH 739

Qy        467 AQWSSGK- SFLLK-KLELSMKTPASDLTEPLPQESWL  IVPLILLDAFHLLLVFAFP  520

Db        740 --YGNIKIVNFLLQHSAKVNAKTKDGY  TPLHQAAQQGHTDIINVLLKNN    - ASPN 790

Qy        521   -VDTNLAIAISLSFLALIYIFSIVIYPGMRSSNSWNWAWALSTRLAPHIGVLELLPK 577

Db        791 ELTVNGNTALGI------- ---- --- --                ARRLGYISVVDT 834

Qy        578 LMPV         -          NPPE       -LPSQTTKADIVKPLPIDYNPLSSV-GGET 613

Db        835 LKIVIEETMTTTTVTEKHKMNVPEIMNEVLDMSTTEVRKANAFEMLSDSEYIGVEEGED 894

Qy        614 SLA---------EMIATLSPACRPFSPIAIPL     --EPVPRIEESCSKKPWRKTCSL 658

Db        895 AMTGDTDKYGGPQDLKELGDDSLIKEGYMGPSLGAPSASLPSPASDRPSYTLN  RSSYAR 952

Qy        659 PSPVIFLFIVSCIIAGITLSIAIFFIIF PIPLI       -VNAILSIAAVVILAFVLNC 709

Db        953 DSMMIEELLVPSKEQHLTPTPEFLGNGL-HYSWAAETLDNVILVSSPILSFLTSFMVTA 992

Qy        710 RTWWQVIFSLLNSQPKRLIHSAASP HELA EGPMKVL-   KIEVEIMAPVAFTIDSPTQ 765

Db        993 RG         GSMRGSPHHG             MKIIFPFPCTPI         -    TR 1019

Qy        766 NQTPLVVIIDGLDAIEQDKVLQMI IVPVLPPA FIAIPAAGFWIIIFAINQNLNSVLP 825

Db       1020 ITCALV         KRHKLAN                 FPDI    GEPPGISSKIV 1067

Qy        826 DSNINGHGYMRN-IVHLPVPLNSP GGNARRELVTLAINGDTTVSDTTVTQELTDRRVGQ 884

Db       1068 EMGPAGAQPLGPVIVSIPHPGSMP H PERRIIVLPSSNGE     --TWRPHQFD - SK 1096

Qy        885 NSLGEMIKL--GSKTALNERDTYAPQVQRTIIRIMNRDLTPELVTEDWPLDISPQCMER 942

Db       1097 NE  DIIELINGMPEELDSTEELPPRIPIIIPIP       -  PQTEAVVA -RIPQ 1142

Qy        943 LLNIGGTGRILRIN     -                                        997

Db       1143 ENNG                                                       1189

Qy        998 TLETN          - -      -P  RL                  IPLVRPIIRI 1042

Db       1190 VKKILGNKATFSPIVTVFIRRAKHHKIIMIIRTISSRGLGNYKGIITNNLELLESI 1249

Qy       1043 DG- - - ---DIRNFEVFLNSRTFVLVAHDVPIRLF-TVDLIIPIREIIADVPAAREQIN 1074

Db       1250 TGGTNFAQWEDI- -----DGITFLTFIKDTVP  -ITNT  RAAFWLAGHQVLETVG 1296

Qy       1075 IGSLAYIPLPLHEGPRPPSGYSQIADVSSAV-NIPFIGVVNGPHGPYYSPELSGPQH 1134

Db       1297 LATQLYFEL----  --- -----   ITV    -- FTYAPFT       VFAKMNLP   1323

Qy       1135 PFYNRAAVPATGSSLLLSSMTVDVVIEPIQIRIIQNRVIQYTITFKANINGRVLSQC 1194

Db       1324 - --------VESSLRCFCMTDDKYDKTLPQGR                 ------ 1346

Qy       1195 NIDELKKEMAMNFGDWHLFRSMVIRMRFVESQVPENPRFIQINSSSAFVHHGE-----S 1248

Db       1347 NFEEV-------------      ARARDIEVLRGKPIYVDNYDNLAFLTKGGQQLVFN 1386

Qy       1249 ARRSSHCELP--- LTELSGQTPY LNPSFREELNTNGLDEGAFNHSNLSNAQSQTRRTPSL 1304

Db       1387 FYSFKENRLPFSIKIRDTSQEPCGRLSFLKERKTTKGLPQTAVCNLNITLPAHKKETES- 1445

Qy       1305 SSLNSQDSSIFISKLTDKVQAEYRDAYRE---YIAQMSQLEGGTGSSTISGRSSPHSTYY 1361

Db       1446 -  DQDDEIE - KTDRRQSFASLALRKRYSYLTEPGMIERSTGAT----RSLP-TTY- 1492

Qy       1362 IGQSSSGGSIHSTLEQERGKEGELKQEDGRKSFLMKRGDVIDYSSSGVSTNEASPLDPIT 1421

Db       1493 -----SYKPFHSTPPYQSWTTAPITVPGPAKS-- ---GFTSLSSSSSNTPSASPLKSIW 1541

Qy       1422 EEDEKSDVSA-KLLPGKKSSERPSLPQTDLKLKGGGLRYQKLPSDEDESGTGRVQITPHC 1481

Db       1542 -- -         --- --SVSTPSPIKSTL----------------GASTTSSVKSISDV 1587

Qy       1482 SKMIRTKRLKAKQRECASPQEHSAEPIRTPIKAKEY---LSDALLDK----------KDS 1528

Db       1588 ASPIRSLPT -          -  MSSPIKTVVSQSPYNIQVSSGTLARAPAVTEATPLKGL 1644

Qy       1529 SDSGVRSNESAPNHSLHNEAADDSQLEKANLIELEDEGHSGKRGMPHS--  -  -- -- 1576

Db       1645 ASNSTFSSPTAP------VTTAGSLLERSSITMTPPASPKSNINMYSSSLPFKSIITSAA 1668

Qy       1577 --  - LSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPA---------------- 1614

Db       1669 PLISSPLKSVVSPVKSRVDVISSAKITMASS---LSSPVKQMPGHAEVALVNGSISPLKY 1725

Qy       1615 CQKAYNLNRTFSTVTLNNN-TAPTNRANQ----NFDEIEGIRETSQVI--  LRPGPSPN 1665

Db       1726 ASSSTLINGCKATATLQEKISSATNSVSSVVSAATDTVEKVFSTTTAMPFSPLRSYVSAA 1785

Qy       1666 PTAVQNENLKSMAHKFSQRSSYTRLSKDASELHAASSEST 1705

Db       1786 PSAFQSLMTPSAS    - ALYTSL --GSSISATTSSVT 1816
```

RESULT 5
YB23_HUMAN
ID YB23_HUMAN STANDARD; PRT; 768 AA.
AC Q9ULJ7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA1223 (Fragment).
GN KIAA1223.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirosawa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC -!- SIMILARITY: Contains at least 14 ANK repeats.
CC ---------- ---------- ---------- ---------- ---------- ---------- ----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---------- ---------- ---------- ---------- ---------- ---------- ----
DR EMBL; AB033049, BAA86537.1; -.
DR HSSP; P42771; 1DC2.
DR InterPro; IPR002110, ANK.
DR Pfam; PF00023; ank, 14.
DR PRINTS; PR01415; ANKYRIN

```
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT        1     11       ANK 1.
FT   REPEAT       1>     44       ANK 2.
FT   REPEAT       48     [illegible]  ANK 3.
FT   REPEAT       86    118       ANK 4.
FT   REPEAT      119    146       ANK 5.
FT   REPEAT      152    181       ANK 6.
FT   REPEAT      185    214       ANK 7.
FT   REPEAT      218    247       ANK 8.
FT   REPEAT      251    280       ANK 9.
FT   REPEAT      284    313       ANK 10.
FT   REPEAT      317    346       ANK 11.
FT   REPEAT      350    379       ANK 12.
FT   REPEAT      383    412       ANK 13.
FT   REPEAT      416    446       ANK 14.
SQ   SEQUENCE   768 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 448;  DB 1;  Length 768;
  Best Local Similarity   28.6%;  Pred. No. 1.6e-17;
  Matches  125;  Conservative   77;  Mismatches  161;  Indels   24;  Gaps    4;

Qy          24 LLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNL 66
               |:  :|:: :  :| :: ||  :  :|| : |:  :|     :|      :
Db           2 LLQHGAEVNKADNDGRTALIAAAYMGHEEIVELLLSHGAEVNHEDVDGRTALSVAAICVP 61

Qy          67 ASKEGHIHIVEELLKSGASLEHRDMYGWTALMWACYKGRTDVVELLLSHGANPSVTGLQY 127
               |:| ||  :|  |:  || :  |   | |:  :  |  :|  |  :|  |  |  |
Db          62 ASKDGHASVVSLLLDRGAEVDHCDPTGMTPLIVAAYEGHVDVVTLLLEHGADVDHTD NN 119

Qy         128 SVYPIIWAAGRGHADIVHLLL                                    158
               :| |  |  | :|:
Db         120 GRTPLLAAASMGHASVVNTLIFWGAAVFSTDP DPTYNGLTALAVCNVEVVKILLDRGLDP 179

Qy         159           ------------ --   QNDAKVVDGDRYSTPLYWAAREGHLEGVKHLL 191
                                           :||:   :  |     |  |:::|| :|:  :
Db         180 NHRDDAGWTPLHYAAFEGHRLIVRALLEQGAPCNEIDNLGRIPLHLAAQEGHYEVVQTLL 239

Qy         192 AMGADVDCEGANSMTALIVAVKQGYTQSVKEILKRNPNVVNLTDKFGNTALMIASKEGHIE 251
                :::|  :  :  | ||  :|  :  :  :  :  :  :|     | :| :  :
Db         240 RNRSNIDQPGYDGRNALRVAALEGHEDIVELLESHGADVNCPEAGGRTPLYILAIENQLT 299

Qy         252 IVQDLIDAGTYVNIPDRSGNTVLIAAVRGDVEIVPALLKYGADIEIRNKTALYWAV 311
               :  :|  |  |  :|  |  |  |  |  |  |  |  |  |  |  |  |  |
Db         300 MAEYLFNHANVEAVDAEGHTALHIAIVNVDIVSTLLKNGALQRAA 460

Qy         312 EKGNATMVKTLL [illegible] 387
               ||:  :  :
Db         461 WCGHVKVVQLLIEHGAVVDATVNDGTSLLHLAAKNGHVECMQVL 419

Qy         472 PLHVAIRGRSRRLAELL 388
                 :|| :   :: :|:
Db         420 AVRVAAKNGHSQTIKLL 436

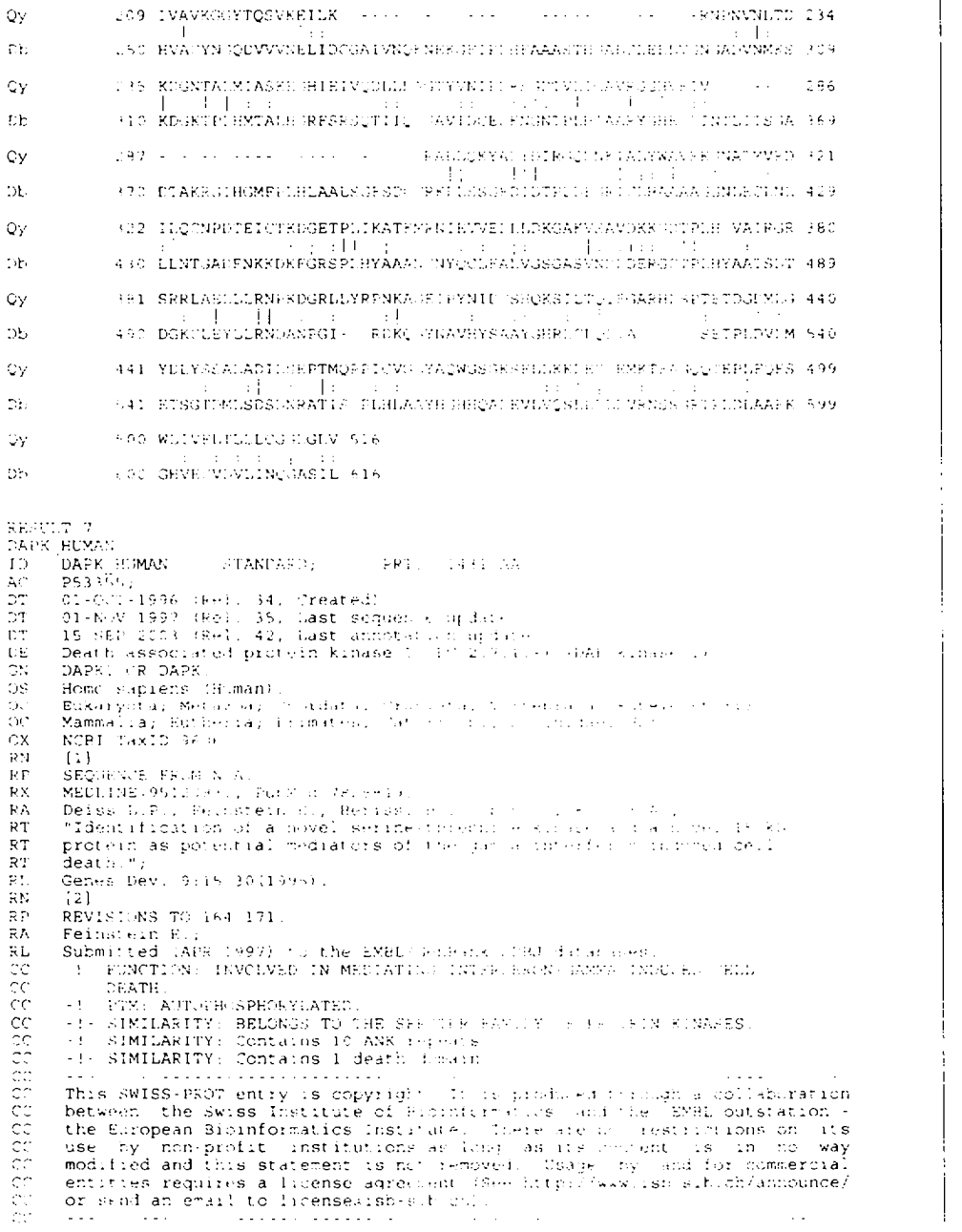

RESULT 6
Y379_HUMAN
ID   Y379_HUMAN     STANDARD;      PRT;  1059 AA.
AC   O15084;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein KIAA0379 (Fragment).
GN   KIAA0379.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97349984; PubMed=9205841;
RA   Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VII.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 4:141-150(1997).
CC   -!- SIMILARITY: Contains at least 27 ANK repeats.
CC   -------- . ---- .. .. ... . ------------------------------ .
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------- -------- . . ------ . . . . ----- --- . .. ------ --
DR   EMBL; AB002377; BAA20833.2; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 28.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 28.
DR   PROSITE; PS50088; ANK_REPEAT; 24.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT       46     75       ANK 1.
FT   REPEAT       79    108       ANK 2.
FT   REPEAT      112    141       ANK 3.
FT   REPEAT      145    174       ANK 4.
FT   REPEAT      178    207       ANK 5.
FT   REPEAT      211    240       ANK 6.
FT   REPEAT      244    273       ANK 7.
FT   REPEAT      277    307       ANK 8.
FT   REPEAT      311    340       ANK 9.
FT   REPEAT      344    373       ANK 10.
FT   REPEAT      377    406       ANK 11.
FT   REPEAT      410    439       ANK 12.
FT   REPEAT      443    472       ANK 13.
FT   REPEAT      476    506       ANK 14.
FT   REPEAT      510    540       ANK 15.
FT   REPEAT      555    584       ANK 16.
FT   REPEAT      588    617       ANK 17.
FT   REPEAT      622    651       ANK 18.
FT   REPEAT      658    687       ANK 19.
FT   REPEAT      691    720       ANK 20.
FT   REPEAT      724    753       ANK 21.
FT   REPEAT      761    790       ANK 22.
FT   REPEAT      793    822       ANK 23.
FT   REPEAT      828    857       ANK 24.
FT   REPEAT      861    891       ANK 25.
FT   REPEAT      895    924       ANK 26.
FT   REPEAT      931    960       ANK 27.
SQ   SEQUENCE   1059 AA;  113465 MW;  C1F55E6CFE494770 CRC64;

  Query Match             4.9%;  Score 431.5;  DB 1;  Length 1059;
  Best Local Similarity   25.1%;  Pred. No. 2.3e-15;
  Matches  155;  Conservative   96;  Mismatches  239;  Indels  127;  Gaps   13;

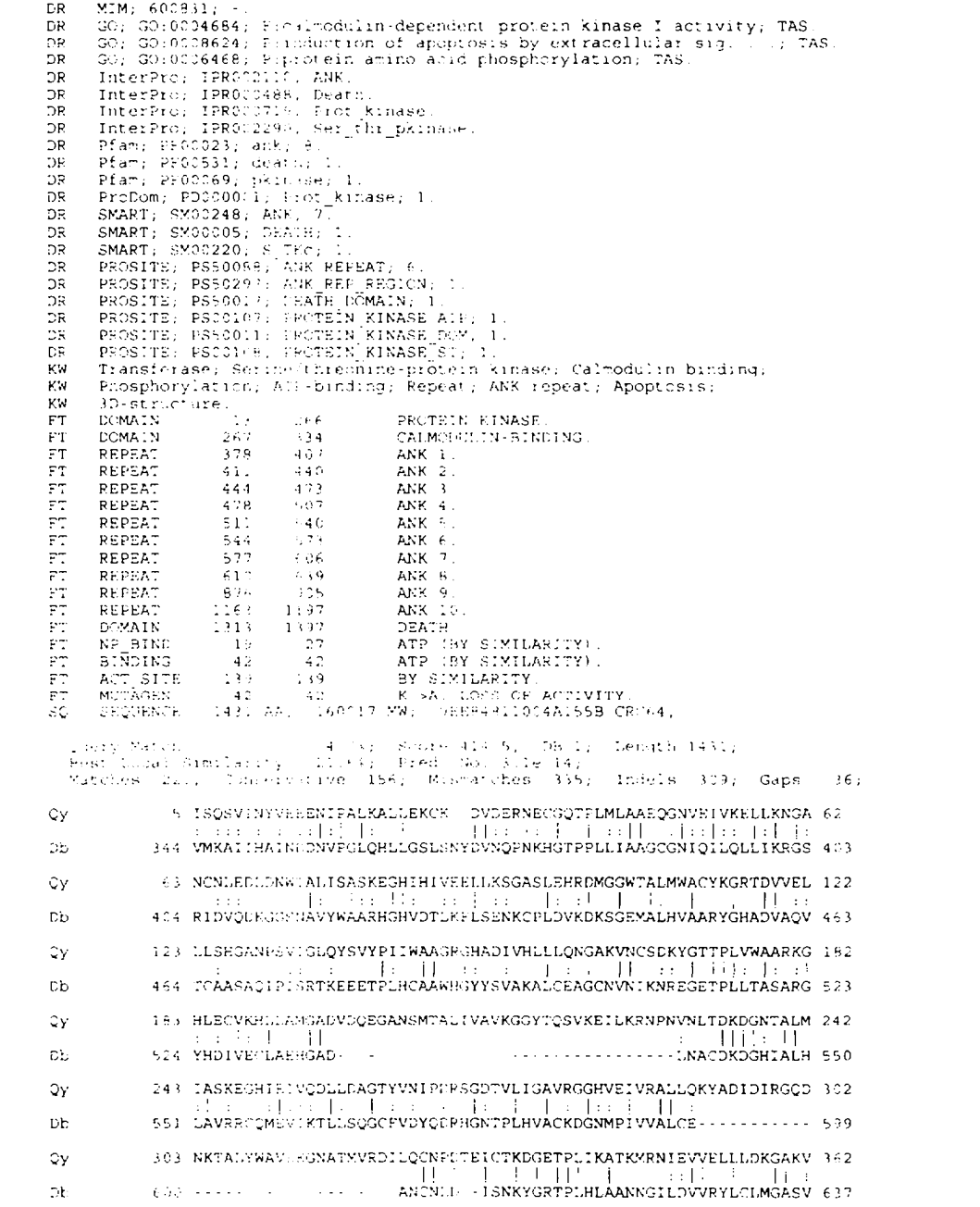
Qy           7 QSVINYVEERNIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNL 66
               |:: :   ::    ::||: : :|:: ::   :||:  ||  |: ||:: :: :||  :
Db          20 QAIFNGDPDE---VRALIFKKEDVNFQDNEKRTPLHAAAYLGDAEIIELLILSGARVNA 75

Qy          67 EDLDNW-----TALISASKEGHIHIVEELLKSGASLEHR---------------------- 100
               :|  |       |: : |::     !: |||  | :  |
Db          76 KD-SKWLTPLHRAVASCSEEA----VQVLLKHSADVNARDKNWQTPLHIAAANKAVKCAE 130

Qy         101 ---------- -DMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGR 148
                         |  | |||  | :  .  ::|:|||| ||| :    :     | |||
Db         131 ALVPLLSNVNVSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFD-KKDRRAIHWAAYM 189
```



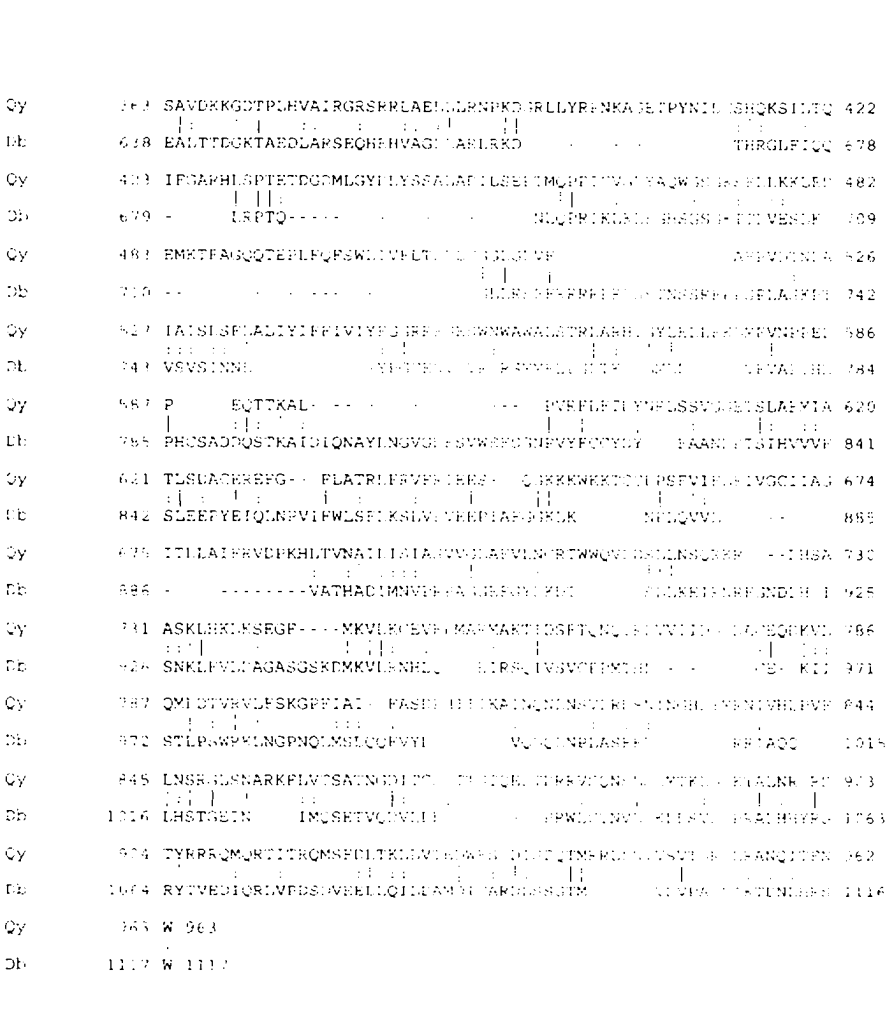
RESULT 7
DAPK_HUMAN
ID DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV 1997 (Rel. 35, Last sequence update)
DT 15 SEP 2003 (Rel. 42, Last annotation update)
DE Death associated protein kinase 1 (EC 2.7.1.-) (DAP kinase 1).
GN DAPK1 OR DAPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96122447; PubMed=7628451;
RA Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT "Identification of a novel serine/threonine kinase and a novel 15-kD
RT protein as potential mediators of the gamma interferon-induced cell
RT death.";
RL Genes Dev. 9:15-30(1995).
RN [2]
RP REVISIONS TO 164-171.
RA Feinstein E.;
RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA INDUCED CELL
CC DEATH.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR PROTEIN KINASES.
CC -!- SIMILARITY: Contains 10 ANK repeats.
CC -!- SIMILARITY: Contains 1 death domain.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---



DR EMBL; X76104; CAA53712.1; -.
DR PDB; 1JKT; 01-APR-02.
DR Genew; HGNC:2674; DAPK1.
DR MIM; 600831; -.
DR GO; GO:0004684; F:calmodulin-dependent protein kinase I activity; TAS.
DR GO; GO:0008624; P:induction of apoptosis by extracellular sig...; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00023; ank; 8.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 7.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50088; ANK_REPEAT; 6.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW 3D-structure.
FT DOMAIN 1 266 PROTEIN KINASE.
FT DOMAIN 267 334 CALMODULIN-BINDING.
FT REPEAT 378 407 ANK 1.
FT REPEAT 411 440 ANK 2.
FT REPEAT 444 473 ANK 3.
FT REPEAT 478 507 ANK 4.
FT REPEAT 511 540 ANK 5.
FT REPEAT 544 573 ANK 6.
FT REPEAT 577 606 ANK 7.
FT REPEAT 617 649 ANK 8.
FT REPEAT 874 905 ANK 9.
FT REPEAT 1164 1197 ANK 10.
FT DOMAIN 1313 1397 DEATH.
FT NP_BIND 10 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 139 139 BY SIMILARITY.
FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
SQ SEQUENCE 1431 AA; 160017 MW; [illegible] CRC64;

Query Match 4.7%; Score 418.5; DB 1; Length 1431;
Best Local Similarity 21.4%; Pred. No. 3.1e-14;
Matches 244; Conservative 156; Mismatches 355; Indels 309; Gaps 36;

Qy 5 ISQSVEINYVRSENIPALKALLEKCR DVDERNECGQTPLMLAAEQGNVEIVKELLKNGA 62
Db 344 VMKAIHAINCDNVPGLQHLLGSLSNYDVNQPNKHGTPPLLIAAGCGNIQILQLLIKRGS 403
Qy 63 NCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVVEL 122
Db 404 RIDVQLKGGNAVYWAARHGHVDTLKFLSENKCPLDVKDKSGEMALHVAARYGHADVAQV 463
Qy 123 LLSHGANPSVIGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKG 182
Db 464 IQAASAQIPISRTKEEETPLHCAAWNGYYSVAKALCEAGCNVNIKNREGETPLLTASARG 523
Qy 183 HLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTALM 242
Db 524 YHDIVEVLAEHGAD- - ----------DNACDKDGHIALH 550
Qy 243 IASKEGHIEIVQDLLDAGTYVNIPDPSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQD 302
Db 551 LAVRRTQMEVIKTLLSQGGCPVDYQDPHGNTPLHVACKDGNMPIVVALCE---------- 599
Qy 303 NKTALYWAVHGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGAKV 362
Db 600 ----- - --- - ANCNIE -ISNKYGRTPLHLAANNGILDVVRYLCLMGASV 637



```
RESULT 8
TNK1_HUMAN
ID   TNK1_HUMAN     STANDARD;      PRT;
AC   O95271; Q95471;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase I) (TNKS) (TNKS1)
DE   Interacting ankyrin-related ADP ribose polymerase
GN   TNKS OR TNKS1 OR TIN1 OR TINF1 OR PARP5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Testis;
RX   MEDLINE=99040105; PubMed=9822378;
RA   Smith S., Giriat I., Schmitt A., de Lange T.;
RT   "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
RL   Science 282:1484-1487(1998).
RN   [2]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=99454782; PubMed=10523501;
RA   Smith S., de Lange T.;
RT   "Cell cycle dependent localization of the telomeric PARP, tankyrase,
RT   to nuclear pore complexes and centrosomes.";
RL   J. Cell Sci. 112:3649-3656(1999).
RN   [3]
RP   FUNCTION, AND PHOSPHORYLATION.
RX   MEDLINE=20556282; PubMed=10988293;
RA   Chi N.-W., Lodish H.F.;
RT   "Tankyrase is a golgi-associated mitogen-activated protein kinase
RT   substrate that interacts with IRAP in GLUT4 vesicles.";
RL   J. Biol. Chem. 275:38437-38444(2000).
RN   [4]
RP   FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.",
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC       subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC       nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
CC       the cytoplasmic domain of LNPEP/Otase in SLC2A4/GLUT4 vesicles.
CC       Binds to the N terminus of telomeric TRF1 via the ANK repeats.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4 vesicles. A minor proportion is
CC       also found at nuclear pore complexes and around the pericentriolar
CC       matrix of mitotic centromeres. During interphase, a small fraction
CC       of TNKS is found in the nucleus, associated with TRF1.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=O95271-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=O95271-2; Sequence=VSP_004538, VSP_004539;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC   -!- PTM: Upon insulin stimulation, phosphorylated on serine residues
CC       by MAPK kinases.
CC   -!- PTM: ADP-ribosylated (auto).
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: C terminus is 85% ANK repeats.
CC   -!- SIMILARITY: C terminus 1 sterile alpha motif (SAM) domain.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF082556; AAC79841.1; -.
DR   EMBL; AF082557; AAC79842.1; -.
DR   EMBL; AF082558; AAC79843.1; -.
DR   EMBL; AF082559; AAC79844.1; -.
DR   HSSP; Q00420; 1AWN.
DR   Genew; HGNC:11941; TNKS.
DR   MIM; 603303; -.
DR   GO; GO:0005696; C:telomere; IDA.
DR   GO; GO:0003950; F:NAD+ ADP-ribosyltransferase activity; IDA.
DR   GO; GO:0005515; F:protein binding activity; IPI.
DR   GO; GO:0007004; P:telomerase-dependent telomere maintenance; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 19.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 17.
```

```
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; ..
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; [illegible]
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
KW   Phosphorylation; Alternative splicing.
FT   REPEAT      215    247       ANK 1.
FT   REPEAT      248    280       ANK 2.
FT   REPEAT      281    313       ANK 3.
FT   REPEAT      368    400       ANK 4.
FT   REPEAT      401    433       ANK 5.
FT   REPEAT      434    466       ANK 6.
FT   REPEAT      521    556       ANK 7.
FT   REPEAT      557    589       ANK 8.
FT   REPEAT      590    622       ANK 9.
FT   REPEAT      683    715       ANK 10.
FT   REPEAT      716    748       ANK 11.
FT   REPEAT      749    781       ANK 12.
FT   REPEAT      846    868       ANK 13.
FT   REPEAT      869    901       ANK 14.
FT   REPEAT      902    934       ANK 15.
FT   DOMAIN     1010   1089       SAM.
FT   DOMAIN     1176   1327       [illegible]
FT   DOMAIN        9     14       [illegible]
FT   DOMAIN       27     34       [illegible]
FT   DOMAIN      108    134       [illegible]
FT   DOMAIN      137    145       [illegible]
FT   VARSPLIC    641    643       PST [illegible]
FT                                /FTId=VSP_004448.
FT   VARSPLIC    644   1347       M[illegible]
FT                                /FTId=VSP_004449.
FT   MUTAGEN    1144   1184       H [illegible] ACTIVITY. WORK [illegible] WITH A.
FT                                WITH A.
FT   MUTAGEN    1201   1231       E->A [illegible] ACTIVITY. WORK [illegible] WITH A.
FT                                WITH A.
SQ   SEQUENCE   1347 AA;  [illegible] MW;  [illegible]

  Query Match             4.5%;  Score [illegible]
  Best Local Similarity   23.3%;  Pred. No. [illegible]
  Matches  167;  Conservative  [illegible]  Mismatches  [illegible]  Gaps  17;

Qy          14 EEENIPALKALLEKKDVIERNRY [illegible] 73
               ||:: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         448 EEKLVALLTPLNVNIHASLGRE [illegible] 494

Qy          74 ALISGFRSDNIHCVRLLEAGA [illegible] 123
               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         495 PIRDGFTPLHVAA [illegible] 544

Qy         124 - [illegible] 124

Db         545 NPDSRSKAVDMAITIGNAEIFEV [illegible] 594

Qy         125 GLQYSV [illegible] 164
               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         595 ALHCAVASLHPKRKQVTELLLKKGTDVNAKTMMGYTPLYMAVE [illegible] 644

Qy         165 NCSDFVSITTPLVWAAAKGHLEVVR [illegible] 207
               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         645 NALDTIAITALRRAALAGHQTCR [illegible] 694

Qy         208 --  --  --LIVAVKSGYTQNVKE [illegible] 256
                       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         695 IRTSDVDYRLLEAAKAGDLETVK [illegible] 743

Qy         257 LDAGTYVNIPDRSGDTVLHDAVRG [illegible] 314
               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         744 LHHGADVHAKDKGGLVPLHNACSY [illegible] 793

Qy         315 - [illegible] 325

Db         794 ELCKILLKHGADPTKKNRQGKTP [illegible] 824
```

```
Qy         326 NPDTEICTKD --GETPLIKATKMRNIEVVELLDKGAKVSAVDKKGDTPLH-------- 374
               |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db         824 TPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHV 883

Qy         375 ---------------- --- - - VAIRGRSRRLAELLLRNPKDGRILLYRPNKAGETP 408
                                       :|||: :| :| : | : :    |: |:.
Db         884 DIAALLIPYNTCVNATDKWAFTPLHEAAQKGRT-QLCALLLAHGADPTM-- KNQEGQTP 939

Qy         409 YN----       -IDPSHQKSILT-- - --QIPGARHLSPTETDGDMLGYDLYSSALADI 461
                :    :     |    :::|    :  :  :  |    |       :::
Db         940 LDLATADIFALLIDAMPPEALNTCFKPCATVVSASLISPASTPS-----CLSAASSIDN 994

Qy         462 LSEPTYQ[illegible] VGLYAQWGSS--  -           -    ----KSFLLKKLDEEMKT 486
               :    :  : ||  .  |                                |:  |:      :|
Db         995 LTGPLAE  LAVGGASNAGIPGAASTERKEGEVAGLDMNISQFLKSLGLEHLRDIFET 1049


RESULT 9
ASB2_HUMAN
ID   ASB2_HUMAN     STANDARD;      PRT;   587 AA.
AC   Q96Q27; Q9NSU5; Q9Y5e7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 2 (ASB-2).
GN   ASB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND INDUCTION.
RX   MEDLINE=21450615; PubMed-11566190;
RA   Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,
RA   Nakanishi T., Tanaka K.;
RT   "ATRA-regulated Asb 2 gene induced in differentiation of HL-60
RT   leukemia cells.";
RL   FEBS Lett. 505:223-228(2001).
RN   [2]
RP   SEQUENCE FROM N.A., AND INDUCTION.
RX   MEDLINE=21634870; PubMed-11682444;
RA   Guibal F.C., Moch-Lutz C., Smolewski P., Di Gioia Y.,
RA   Darzynkiewicz Z., Lutz P.G., Cayre Y.E.;
RT   "ASB-2 inhibits growth and promotes commitment in myeloid leukemia
RT   cells.";
RL   J. Biol. Chem. [illegible]:18224-18232(2002).
RN   [3]
RP   SEQUENCE OF 42-[illegible] FROM N.A.
RC   TISSUE=Brain;
RA   Ohlenschlaeger B., Heimann B., Meyer H.W., Hassenhuber C.,
RA   Wiemann S.;
RL   Submitted JAN 2[illegible] to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 42-[illegible] FROM N.A.
RX   MEDLINE=20564128; PubMed-11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb 1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2.00).
CC   -!- FUNCTION: Not known. Its expression induces growth inhibition and
CC       chromatin condensation.
CC   -!- INDUCTION: By all-trans retinoic acid (ATRA).
CC   -!- SIMILARITY: Contains 12 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AB056723; BAB64532.1; -.
DR   EMBL; AJ251238; CAC17765.1; -.
DR   EMBL; AL137735; CAB70899.1; -.
DR   EMBL; AF159164; AAD45345.1; -.
DR   PIR; T46507; T46507.
DR   Genew; HGNC:16012; ASB2.
DR   MIM; 605759; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS.
DR   Pfam; PF00023; ank; 11.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       56     85       ANK 1.
FT   REPEAT       89    119       ANK 2.
FT   REPEAT      123    152       ANK 3.
FT   REPEAT      156    185       ANK 4.
FT   REPEAT      189    218       ANK 5.
FT   REPEAT      222    251       ANK 6.
FT   REPEAT      255    284       ANK 7.
FT   REPEAT      288    317       ANK 8.
FT   REPEAT      320    349       ANK 9.
FT   REPEAT      362    391       ANK 10.
FT   REPEAT      392    421       ANK 11.
FT   REPEAT      428    456       ANK 12.
FT   DOMAIN      533    587       SOCS BOX.
FT   CONFLICT    224    224       MISSING (IN REF. 4).
FT   CONFLICT    460    460       MISSING (IN REF. 4).
SQ   SEQUENCE   587 AA;  65083 MW;  86A3B96010C4C1FD CRC64;

  Query Match             4.3%;  Score 484.5;  DB 1;  Length 587;
  Best Local Similarity   30.4%;  Pred. No. 3.1e-13;
  Matches  113;  Conservative   71;  Mismatches  161;  Indels   27;  Gaps    6;

Qy          9 VINYVEEENIPALKALLEKCKDVDERNEVGQTPLMLAAEQGNVEIVKELLKNGANCNLE 67
Db         61 LIKAIKDGDEEALKTMIKEGKNLAEPNKEGWLPLHEAAYYGQVGCLKVLQRAYPGTIDQR 120

Qy         68 DLDNWTALISASKEGHIHIVEELLKCGASLEHRDMGGWTALMWACYKGRTDVVELLLSHG 127
Db        121 TLQEETAVYLATCRGHLDCLLSLLQAGAEPDISNKSRETPLYKACERKNAEAVKILVQHN 180

Qy        128 A-       NPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAK       VNCSDKYGTTPLVWAARK 181
Db        181 ADTNHRCNRGWTALHESV                  ...             ... 233

Qy        182 GHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTAL 241
Db        234 GQLEALRFLAKYGADINTQASDNAAALYEACKNEHEEVVEFLLSQGADANKTNKDGLLPL 293

Qy        242 MIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADID---- 297
Db        294 HIASKKGNYRIVQMLLPVTSRTRIRRSGVSPLHLAAERNHDEVLEALLSARFDVNTPLA 352

Qy        298 ----IRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVV 352
Db        353 PERARLYEDRRSSALYFAVVNNNVYATELLLQHGADP   NRDVISPLLVAIRHGCLRTM 409

Qy        353 ELLLDKGAKVSA 364
Db        410 QLLLDHGANIDA 421


RESULT 10
TNK2_HUMAN
ID   TNK2_HUMAN     STANDARD;      PRT;  1166 AA.
AC   Q9H2K2; Q9H8F2; Q9HAS4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-
DE   interacting ankyrin related ADP-ribose polymerase 2) (Tankyrase like
DE   protein) (Tankyrase related protein).
GN   TNKS2 OR TNKL OR TANK2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=21072550; PubMed=11205898;
RA   Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W.,
RA   Glass B., Meese E.U.;
RT   "Novel tankyrase-related gene detected with meningioma-specific
RT   sera.";
RL   Clin. Cancer Res. 7:113-119(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast carcinoma;
RX   MEDLINE=21190090; PubMed=11294570;
RA   Kuimov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,
RA   Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;
RT   "Cloning and characterization of TNKL, a member of tankyrase gene
RT   family.";
RL   Genes Immun. 2:52-55(2001).
RN   [3]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC   TISSUE=Liver;
RX   MEDLINE=21264473; PubMed=11278563;
RA   Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,
RA   Eyre H.J., Sutherland G.R., Daly R.J.;
RT   "Identification of a novel human tankyrase through its interaction
RT   with the adaptor protein Grb14.";
RL   J. Biol. Chem. 276:17172-17180(2001).
RN   [4]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC   TISSUE=Placenta;
RX   MEDLINE=21443728; PubMed=11454873;
RA   Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D.,
RA   Morin G.B., Yaswen P., Campisi J.;
RT   "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
RT   rapid induction of cell death upon overexpression.";
RL   J. Biol. Chem. 276:35891-35899(2001).
RN   [5]
RP   SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND
RP   LNPEP/OTASE.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=21661461; PubMed=11802774;
RA   Sbodio J.I., Lodish H.F., Chi N. W.;
RT   "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1
RT   (telomere-repeat-binding factor 1) and IRAP (insulin-responsive
RT   aminopeptidase).";
RL   Biochem. J. 361:451-459(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Yin Y., Gelmann E.P.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Tracey A.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 838-1166 FROM N.A.
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
```

```
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   FUNCTION, AND ADP-RIBOSYLATION.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-R.. ...) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC    !- FUNCTION: May regulate vesicle trafficking and ... dulate the
CC       subcellular distribution of ... A4/GLUT4 vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (A...P-ribosyl)(n) = nicotinamide + (ADP-D-ribosyl)(n+1) acceptor.
CC   -!- SUBUNIT: Oligomerizes and ass...ates with TNKS. Interacts with the
CC       cytoplasmic domain of LNPEP/I...ase in GLUT2A4/GLUT4 vesicles. Binds
CC       to the N-terminus of GRB14 and TRF1 with its ankyrin repeat
CC       region.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4 vesicles. Also found at and the
CC       pericentriolar matrix of mitotic centromeres during interphase, a
CC       small fraction of TNKS2 is found in the nucleus, associated with
CC       TRF1.
CC   -!- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,
CC       liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral
CC       blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,
CC       ovary, small intestine, colon, mammary gland, breast and breast
CC       carcinoma, and in common-type meningioma. Highly expressed in
CC       fetal liver, heart and brain.
CC   -!- PTM: ADP-ribosylated (auto).
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC    !- SIMILARITY: Contains 25 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF305961; AAG25674.1; ALT_INIT.
DR   EMBL; AF264912; AAG44634.1; -.
DR   EMBL; AF329695; AAK21468.1; -.
DR   EMBL; AF342982; AAK25811.1; -.
DR   EMBL; AF304831; AAK81...1; -.
DR   EMBL; AF439...; AAL4...1; -.
DR   EMBL; AL...; ...; -.
DR   EMBL; AK02...; BAB14...1; -.
DR   HSSP; Q00...; 1AWC.
DR   Genew; HGNC:...; TNKS2.
DR   MIM; 607128; -.
DR   InterPro; IPR001110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 18.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 25.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.
FT   REPEAT       57     89       ANK 1.
FT   REPEAT       90    122       ANK 2.
FT   REPEAT      123    155       ANK 3.
FT   REPEAT      210    242       ANK 4.
FT   REPEAT      243    275       ANK 5.
FT   REPEAT      276    308       ANK 6.
FT   REPEAT      363    398       ANK 7.
FT   REPEAT      399    431       ANK 8.
FT   REPEAT      432    464       ANK 9.
FT   REPEAT      525    557       ANK 10.
FT   REPEAT      558    590       ANK 11.
FT   REPEAT      591    623       ANK 12.
FT   REPEAT      678    710       ANK 13.
FT   REPEAT      711    743       ANK 14.
FT   REPEAT      744    776       ANK 15.
FT   DOMAIN      870    936       SAM.
FT   DOMAIN     1026   1162       PARP.
FT   CONFLICT     331    337       KGHSLLQ -> QRPLVAA (IN REF. 1)
FT   CONFLICT     359    361       NFKHP -> IQAS (IN REF. 1).
FT   CONFLICT     966    966       Q -> P (IN REF. 8).
SQ   SEQUENCE   1166 AA;  126217 MW;  4F8B5B8D97CAF704 CRC64;

  Query Match             4.3%;  Score 380.5;  DB 1;  Length 1166;
  Best Local Similarity   24.0%;  Pred. No. 1.5e-12;
  Matches  167;  Conservative   73;  Mismatches  206;  Indels  217;  Gaps   13;

Qy          14 EERNIPAIHALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWT 73
               |:: :    |   |   ' |    || |||    |:||: ||:: : : :  :
Db         190 EEKMMMALIPLNVNCHASDGRK- -STPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy          74 ALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVVELLLSHGANPSVT 133
                : :|   ||  : |  :| |:    :  :|    |  ' : :|   |'|: :|::
Db         247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLD 306

Qy         134 G-----        ---------LQYS--  ---------------------     ----- -- 138
                             |:
Db         307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy         139 -------     ----------------     ----VYPIIWAAGRGHADIVHLLLQNGAKV 164
                                                : |:  : : | :: | ::::: |||
Db         367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy         165 NGSDKYGITPLVWAAPKGHLECVKHLLAMGADVD----------------  --QEG-- 201
                  |    |   ||   |:  :  ||: : | :                        :.|
Db         427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy         202 ---ANSMTALIVAVKGRYTQSVKE-        : ----------  :  ----  -----IL 224
                  : :  |.  : |     :: ::                              :.
Db         487 LGNSEALKQLDEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy         225 KRNPNVNIRTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVE 284
               :  :  :  |  |     |  .:| :  :  |:  :  |:  |       |     | ,
Db         547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYR 606

Qy         285 INPALIQKVAEDIRSQDNKTALYWAVEKGNATVVRDLLQ----  -  -          324
                :  :   :  :  |     |    :  :|:
Db         607 ...KDL...GWETTKKNRDGNTPL...DLVKDGDTLLQDLLERGDAALLDAAKKGCLARVKKL 664

Qy         325 CNPDTRDHEGDR--  -TPLIKATKMRNIEVVEILLKKGAKVSAVDKKGDTPLH------ 374
               :||  : :  :    ||| :   :||  ||   :  :|  :  | :,|
Db         665 SSPDNVDKRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 723

Qy         375 -------     ---------------- VAIRGRSRRLAELLLRNPKDGRLLYRPNKAGE 406
                                            | : |: :   |||  :  |   :  |: ::
Db         724 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRT-QLCALLLAHGADPTL---KNQEGQ 779

Qy         407 TPYNIDQNHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQP 459
               | ::  :  | :           :     :|       : || ||     :
Db         780 TPLDLVSAIDVSALLTAAMPPSALPSCYKPQVINGVPSPGATADALSSGPSSP 832


RESULT 11
ANKH_CHRVI
ID   ANKH_CHRVI     STANDARD;      PRT;   323 AA.
AC   Q06527;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ankyrin homolog precursor.
OS   Chromatium vinosum.
```

```
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC   Chromatiaceae; Allochromatium.
OX   NCBI_TaxID=1049;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D / ATCC 17899 / DSM 180;
RX   MEDLINE=93300843; PubMed=8390993;
RA   Dolata M.M., van Beeumen J.J., Ambler R.P., Meyer T.E.,
RA   Cusanovich M.A.;
RT   "Nucleotide sequence of the heme [illegible] from
RT   the purple phototrophic bacterium, Chromatium vinosum. A [illegible] kilobase
RT   pair DNA fragment contains two cytochrome [illegible] a flavoprotein,
RT   and a homolog of human ankyrin.";
RL   J. Biol. Chem. 268:14426-14431(1993).
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------- [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- [illegible]
DR   EMBL; L13419; AAA23315.1; -.
DR   PIR; B47169; B47169.
DR   HSSP; P55273; 1BD8.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Signal; ANK repeat; Repeat.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    323       ANKYRIN HOMOLOG.
FT   REPEAT       31     60       ANK 1.
FT   REPEAT       64     93       ANK 2.
FT   REPEAT       97    126       ANK 3.
FT   REPEAT      130    159       ANK 4.
FT   REPEAT      163    192       ANK 5.
FT   REPEAT      196    225       ANK 6.
FT   REPEAT      229    258       ANK 7.
FT   REPEAT      262    291       ANK 8.
SQ   SEQUENCE   323 AA;  33459 MW;  [illegible]

  Query Match             4.8%;  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   [illegible]
  Matches   47;  Conservative   [illegible]

Qy          24 DLEKKTVDERNKLQTFLYLAAE[illegible]                    81

Db          61 DLDQGAHVDAPNQFGKTALMFAVE[illegible]                  119

Qy          84 IHIVEELLKSGANLEHRDMGGWTALMWAVYFGHLEVVELLI[illegible] 143

Db         111 IGISALDLERGAHVHIDRTRSGWDALYIAKRHGIIEVVRLLI[illegible] 169

Qy         144 WAAGRGHADIVHLLLQNGAKVNCSKYGITPLYVAAHFGHIEIAMGALVQQGGAN 203

Db         170 QAASKGRTGVLPLLIEGGADLEARKKDATALHKAAFFGKAEACQLLIQHQIDAVDAL 229

Qy         204 SMTALIVAVKGGYTQSVKQILKRNINVNITDKFNRTALMLAVFHGHEEIVLDLIQY 263

Db         230 GSTALILAVGHGHVAMVERLLAMGA[illegible]                   289

Qy         264 NIPDRSGDTVLIGAVRGGHVEIVRA[illegible]                  [illegible]

Db         290 DLKDADRTAADIAVQLGHRQAIAE[illegible]                  [illegible]


RESULT 12
ANR3_HUMAN
ID   ANR3_HUMAN     STANDARD;      PRT;   832 AA.
AC   P57078; Q96KH0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Serine/threonine protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat
DE   domain protein 3) (PKC delta-interacting protein kinase).
GN   ANKRD3 OR DIK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal kidney, and Fetal lung;
RA   Shimizu N., Kudoh J., Shibuya K.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289793; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P57078-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P57078-2; Sequence=VSP_004862;
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   ---------- [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- [illegible]
DR   EMBL; AB047783; BAB56136.1; -.
DR   EMBL; AP001743; BAA95526.1; -.
DR   HSSP; P25963; 1IKN.
DR   Genew; HGNC:496; ANKRD3.
DR   MIM; 606706; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00023; ank; 10.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   ANK repeat; Alternative splicing.
```

```
FT   DOMAIN       22    264       PROTEIN KINASE.
FT   REPEAT      485    514       ANK 1.
FT   REPEAT      518    547       ANK 2.
FT   REPEAT      551    580       ANK 3.
FT   REPEAT      584    613       ANK 4.
FT   REPEAT      617    647       ANK 5.
FT   REPEAT      651    680       ANK 6.
FT   REPEAT      684    713       ANK 7.
FT   REPEAT      717    746       ANK 8.
FT   REPEAT      750    779       ANK 9.
FT   REPEAT      782    811       ANK 10.
FT   NP_BIND      28     36       ATP (POTENTIAL).
FT   BINDING      51     51       ATP (POTENTIAL).
FT   ACT_SITE    148    148       BY SIMILARITY.
FT   VARSPLIC    274    345       Missing (in isoform 2).
FT                                 /FTId=VSP_004442.
FT   CONFLICT    714    714       V -> A (IN REF. 2).
SQ   SEQUENCE   832 AA;  96610 MW;  5C4B8D3BF04F3B7B CRC64;

  Query Match             4.1%;  Score 466;  DB 1;  Length 832;
  Best Local Similarity   37.5%;  Pred. No. 1.1e-11;
  Matches  139;  Conservative   40;  Mismatches  126;  Indels  102,  Gaps    6;

Qy         29 KQVDERNECGQTPLMLAARQGNVEIVKFLLNHGANVNL        ---  -----    - 66
              :|||  : :|  : :|:  :  |    |
Db        477 QDVDGLTESGAAPLSLAVRAEQAEVAEVLLKAGANINSGD-  GETPLHVAAAMGYVDV 535

Qy         67          --EDLDNWTALHWAFLNGHEQVAEVLVQNGVAVNATDRWGNTALHWAAYVNNVEA 123
                       :|    :|:    :    |:   :    :    :  :  :  :  :  :
Db        536 ELLAARGSVNAKDDQWTALHFAAIDD-ESMVELLERNAQIDVQDE-EGMHVAVQHG 590

Qy        124 RDVVDVLLLSHGADPSVTNQYSYTPLHWAATKNRIDIVNLLQKYQASRDLETGDL 178
              : :|:  :   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        591 QENIVSLLQRRGVDVSLQT-RDAKGETVLAAA--RLCVLQSVKKLLAAGTLT-PTF 645

Qy        179 DVWAAAFFGHLEVVEHLLAYGADVNAQDKFGKTAFDISIGNGNEDLAEILQKLN 234
              :  |::|    : :|:  :|:    :  |  :    :  :  :  :  :  :
Db        646 DRLAAQSGRYRVAKLIIETGSDVN-TPDSRGRTALHLAVAAG-YDRIVRLLLSRGADRAM 705

Qy        235 KDGNTALMIASKEGHIEVVQELLKYKADVNAKDKVGLTALHLAAQGGHLEIVEVLLKNGA 294
              :|:  :|:|:  :  |:  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        706 SDGYTALHLAA-  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  - 742

Qy        295 DIDIKNKDNKTALYWAVEFNNMDIVKLLIKYSANVNAVDNHGVTPLDLVKDRMSERLLKA 354
              :  :  :|:|:  :  |  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        743 DVLARDSLNQTALHLAAAIGHSEVVSV-  -  -  -  -  -  -  -  -  SATDVLL 798

Qy        355 LIDFGVNAVTPSGETPLDLAKGLGYEEIVGVLQQAEAKHAQNRLEDKDDEDSEMETA 412
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        799 F-  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  - 832


RESULT 13
RNSA_HUMAN
ID   RNSA_HUMAN     STANDARD;      PRT;   741 AA.
AC   Q05823;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease F).
GN   RNASEL OR RNS4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-394 AND GLU-541.
RC   TISSUE=Kidney;
RX   MEDLINE=93201598; PubMed=7680958;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A dependent RNAse: a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:753-765(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20515260; PubMed=11163265;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes: mediators
RT   of interferon action.";
RL   Mamm. Genome 11:994-997(2000).
RN   [3]
RP   CHARACTERIZATION OF RNASEL ACTIVITY.
RX   MEDLINE=94245727; PubMed=7514601;
RA   Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
RA   Silverman R.H.;
RT   "Intrinsic molecular activities of the interferon-induced 2-5A-
RT   dependent RNase.";
RL   J. Biol. Chem. 269:14153-14158(1994).
RN   [4]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11585831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";
RL   J. Biol. Chem. 276:48473-48482(2001).
RN   [5]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13954-13954(2002).
RN   [6]
RP   REVIEW.
RX   MEDLINE=99073516; PubMed=9856285;
RA   Castelli J., Wood K.A., Youle R.J.;
RT   "The 2-5A system in viral infection and apoptosis.";
RL   Biomed. Pharmacother. 52:386-390(1998).
RN   [7]
RP   MUTAGENESIS OF LYS-394.
RX   MEDLINE=99081837; PubMed=9862963;
RA   Dong B., Silverman R.H.;
RT   "Alternative function of a protein kinase homology domain in 2',
RT   5'-oligoadenylate dependent RNase L.";
RL   Nucleic Acids Res. 27:439-445(1999).
RN   [8]
RP   MUTAGENESIS OF HIS-583; PRO-584, TRP-632; ASP-661; ARG-667 AND
RP   HIS-672.
RX   MEDLINE=21230542; PubMed=11333810;
RA   Dong B., Niwa M., Walter P., Silverman R.H.;
RT   "Basis for regulated RNA cleavage by functional analysis of RNase L
RT   and Ire1p.";
RL   RNA 7:361-373(2001).
RN   [9]
RP   VARIANTS SER-54, LEU-426; GLN-462 AND GLU-541.
RX   MEDLINE=21947464; PubMed=11941534;
RA   Rokman A., Ikonen T., Seppaelae E.H., Nupponen N., Autio V.,
RA   Mononen N., Bailey-Wilson J., Trent J., Carpten J., Matikainen M.P.,
RA   Koivisto P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
RT   "Germline alterations of the RNASEL gene, a candidate HPC1 gene at
RT   1q25, in patients and families with prostate cancer.";
RL   Am. J. Hum. Genet. 70:1299-1304(2002).
RN   [10]
RP   VARIANTS SER-54, GLN-462 AND GLU-541.
RX   MEDLINE=21676852; PubMed=11793394;
RA   Carpten J., Nupponen N., Isaacs S., Sood R., Robbins C., Xu J.,
RA   Faruque M., Moses T., Ewing C., Gillanders E., Hu P., Bujnovszky P.,
RA   Makalowska I., Baffoe-Bonnie A., Faith D., Smith J., Stephan D.,
RA   Wiley K., Brownstein M., Gildea D., Kelly B., Jenkins R.,
RA   Hostetter G., Matikainen M., Schleutker J., Klinger K., Connors T.,
RA   Xiang Y., Wang Z., De Marzo A., Papadopoulos N., Kallioniemi O. P.,
RA   Burk R., Meyers D., Greenberg B., Meltzer P., Silverman R.,
RA   Bailey-Wilson J., Walsh P., Isaacs W., Trent J.;
RT   "Germline mutations in the ribonuclease L gene in families showing
RT   linkage with HPC1.";
```

```
RL   Nat. Genet. 30:181-184(2002).
RN   [11]
RP   CHARACTERIZATION OF VARIANTS GLN-462 AND HIS-541.
RX   MEDLINE=22345000; PubMed=12415269;
RA   Casey G., Neville P.J., Plummer S.J., Xiang Y., Krumroy L.M.,
RA   Klein E.A., Catalona W.J., Nupponen N., Carpten J.D., Trent J.M.,
RA   Silverman R.H., Witte J.S.;
RT   "RNASEL Arg462Gln variant is implicated in up to 13% of prostate
RT   cancer cases.";
RL   Nat. Genet. 32:581-583(2002).
CC   -!- FUNCTION: Endoribonuclease, activated by interferon-A (2'-5'), which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of RNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 2'-5' linked dimers, with preference for
CC       UU and UA sequences, to yield 5'-hydroxyl products ranging from
CC       between 4 and 22 nucleotides in length.
CC   -!- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC       either manganese or magnesium and ATP.
CC   -!- ENZYME REGULATION: After binding to 2-5A (2'-phosphorylated 2',5'
CC       linked oligoadenylates) the homodimerization and subsequent
CC       activation occurs. Inhibited by RNase L inhibitor.
CC   -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC   -!- TISSUE SPECIFICITY: Highly expressed in spleen and thymus followed
CC       by prostate, testis, uterus, small intestine, colon and peripheral
CC       blood leukocytes.
CC   -!- INDUCTION: By interferons.
CC   -!- DOMAIN: The kinase domain allows the homodimerization.
CC   -!- DOMAIN: The nine ankyrin repeats also called 2-5A motif
CC       constitute the 2-5A binding domain.
CC   -!- DISEASE: Defects in RNASEL may be the cause of susceptibility to
CC       hereditary prostate cancer 1 (HPC1) [MIM:601518].
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; D10381; AAA18032.1; -.
DR   HSSP; P80144; 2MYO.
DR   Genew; HGNC:10050; RNASEL.
DR   MIM; 180435; -.
DR   MIM; 601518; -.
DR   GO; GO:0004521; F:endoribonuclease activity; IDA.
DR   GO; GO:0003725; F:RNA binding; IDA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000567; PUG.
DR   Pfam; PF00023; ank; 9.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00580; PUG; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION;
DR   PROSITE; PS50011; PROTEIN_KINASE; 1.
KW   Hydrolase; Nuclease; RNA-binding; Antiviral defense; Ankyrin repeat;
KW   ANK repeat; Polymorphism.
FT   REPEAT       24     53       ANK 1.
FT   REPEAT       58     87       ANK 2.
FT   REPEAT       91    120       ANK 3.
FT   REPEAT      124    153       ANK 4.
FT   REPEAT      167    197       ANK 5.
FT   REPEAT      201    234       ANK 6.
FT   REPEAT      238    268       ANK 7.
FT   REPEAT      272    301       ANK 8.
FT   REPEAT      303    329       ANK 9.
FT   DOMAIN      365    586       PROTEIN KINASE LIKE.
FT   DOMAIN      587    741       RIBONUCLEASE.
FT   DOMAIN      229    242       2-5A BINDING (P-LOOP) 1.
FT   DOMAIN      253    275       2-5A BINDING (P-LOOP) 2.
FT   ZN_FING     398    444       C6 TYPE (ATYPICAL).
FT   VARIANT      59     59       G -> S.
FT                                /FTId=VAR_013519.
FT   VARIANT     406    406       S -> F.
FT                                /FTId=VAR_013520.
FT   VARIANT     462    462       R -> Q (risk factor for prostate cancer;
FT                                reduced enzymatic activity;
FT                                dbSNP:486907).
FT                                /FTId=VAR_012056.
FT   VARIANT     541    541       D -> E (no change in enzymatic activity;
FT                                dbSNP:627928).
FT                                /FTId=VAR_012057.
FT   MUTAGEN     240    240       K ->N: REDUCED 2-5A BINDING ACTIVITY;
FT                                ALMOST COMPLETE LOSS OF 2-5A BINDING
FT                                ACTIVITY WHEN ASSOCIATED WITH N-274.
FT   MUTAGEN     274    274       K->N: REDUCED 2-5A BINDING ACTIVITY;
FT                                ALMOST COMPLETE LOSS OF 2-5A BINDING
FT                                ACTIVITY WHEN ASSOCIATED WITH N-240.
FT   MUTAGEN     392    392       K->R: COMPLETE LOSS OF ENZYMATIC ACTIVITY
FT                                AND ENZYME DIMERIZATION. NO CHANGE IN
FT                                BINDING TO 2-5A AND RNA.
FT   MUTAGEN     583    583       H->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN     584    584       P->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN     632    632       W->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN     661    661       D ->A: COMPLETE LOSS OF ENZYMATIC
FT                                ACTIVITY.
FT   MUTAGEN     667    667       R ->A: COMPLETE LOSS OF ENZYMATIC
FT                                ACTIVITY. NO CHANGE IN 2-5A BINDING AND
FT                                ENZYME DIMERIZATION.
FT   MUTAGEN     672    672       H->A: COMPLETE LOSS OF ENZYMATIC
FT                                ACTIVITY. NO CHANGE IN 2-5A BINDING

  Query Match          4.0%;  Score 357.5;  DB 1;  Length 741;
  Best Local Similarity   27.4%;  Pred. No. 1.2e-11;
  Matches  131;  Conservative   64;  Mismatches  128;  Indels  129;  Gaps   13;

Qy        42 LMLAARQGDVEEVKELLKNGANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRD 101
             :| | :|:|:: ::||: ||| :: :                                :
Db        59 LIKAVQNH VDVVQQLLEGGANVNFQE- ----- --------- --          --E 96

Qy       102 MGGWTALMWACYKGRTDVVELLLSHGANPSVCGLQYSVYPIIWAAGRGHADIVHLLLQNG 161
              :|| | | | | |:- |:| | | :
Db        97 SGGWTPLIWAAVQGMSRDIVELLLPHGADPVL-- - --------- --- ----- -- 87

Qy       162 -- AFVNE- - :::|VWAAFKGLPEVVKFLIAMGADVNQEGANSMIALIVAVAGGYTQSVK 221
                  :|  :  | | :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       88      PFL-VAIRFLAAIAGRVDLIKLPISKGADVNEPLFYGFTAFYEAAVYGKVKALK 142

Qy       222 RILKRNPFL-NL           TDFHSNTALVIASKEGHIEIVQDLLD-AGTYVNIPDRSG 279
             : :|     |            |  |:|' .:::.|:::   | |   :  | |
Db       143 FLYKRGANVDLR-RKTKEDQERLRKGGATALMDAAEKGHVEVLKILLDEMGADVNACDNMG 202

Qy       280 DTVLIGAVRG   GHVE-IVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCN 338
             :| |:       ||  :  |'  ||:::||:  :| |  '|| :  :|: :|
Db       203 RNALIHAL SGDDSDVEAITHLLLDHGADVNVRGERGKTPLILAVEKKHLGLVQRLLEQE 262

Qy       339 -  PDTE-CTKDGETPLIKATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSR 392
                ||:     :|| | | :::  :: '|| :'|        || :  | ||
Db       263 HIEINDTL -SDGKTALILAVELKLKKIAELLCKRGASTDC----GDLVM TARRNYDH 314

Qy       393 RLAELLLRNP   --  -- -------------KDGRLLYRPNKAGETPYNIDCSHQKSI 439
              | ::|| :                        ||  :|||   |: : :|  ::
Db       315 SLVKVLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRP-MIGKLKPFIDEKYK--- 370

Qy       440 LTQIFGAHLSPTETDGDMLGY 461
                  :: |   |  |:
Db       371 ----- - IAVTSEGGIYLGF 393
```

```
RESULT 14
RN_A_MOUSE
ID   RN5A_MOUSE     STANDARD;      PRT;   735 AA.
AC   Q05921; Q9ERU7;
DT   01-NOV 1995 (Rel. 32, Created)
DT   28-FEB 2003 (Rel. 41, Last sequence update)
DT   28-FEB 2003 (Rel. 41, Last annotation update)
DE   2',5'-oligoadenylate-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN   RNASEL OR RNS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3H/An; TISSUE=Adipose tissue;
RX   MEDLINE=20519260; PubMed=11062256;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes and pseudogenes
RT   of interferon action.";
RL   Mamm. Genome 11:989-994(2000).
RN   [2]
RP   SEQUENCE OF 1-679 FROM N.A.
RX   MEDLINE=93201538; PubMed=7681356;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAases: a uniquely regulated mediator
RT   mediator of interferon action.";
RL   Cell 72:753-765(1993).
RN   [3]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11945831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in L929 fibrosarcoma cells.";
RL   J. Biol. Chem. 276:48473-48482(2001).
RN   [4]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13354-13354(2002).
CC   -!- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of mRNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 3' UpNp bonds to 3'-phosphomonoesters for
CC       2C and 2A sequences, to 3'-phosphodiesters for 3C sequences.
CC       between 6 and 22 nucleotides apart.
CC   -!- COFACTOR: Optimal RNA cleavage requires Mg2+ ions.
CC   -!- ENZYME REGULATION: Activated by 2',5'-linked oligoadenylates (2-5A).
CC       Binding of 2-5A to the N-terminal ankyrin repeat domain induces
CC       activation dimers. Inhibited by RNase L inhibitor.
CC   -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasm; and microsomes.
CC   -!- TISSUE SPECIFICITY: Expressed in spleen, thymus, lung, testis,
CC       kidney, liver and heart.
CC   -!- INDUCTION: By interferons.
CC   -!- DOMAIN: A single active nuclease domain is sufficient for
CC       sufficient for ribonuclease activity.
CC   -!- DOMAIN: The kinase domain allows the dimerization.
CC   -!- DOMAIN: The nine ankyrin repeats also inhibit RNase activity;
CC       constitute the 2-5A binding domain.
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   -!- SIMILARITY: Contains 1 ribonuclease domain.
CC   -!- SIMILARITY: Contains 1 zinc-finger C6-type.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to licenses@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF281045; AAG33708.1; -.
DR   EMBL; L10382; AAA37117.1; -.
DR   PIR; B45771; B45771.
DR   HSSP; P42773; 1IHB.
DR   MGD; MGI:1098272; Rnasel.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000907; PUG.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00590; PUG; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger;
KW   Repeat; ANK repeat.
FT   REPEAT      24     53       ANK 1.
FT   REPEAT      58     87       ANK 2.
FT   REPEAT      91    120       ANK 3.
FT   REPEAT     124    153       ANK 4.
FT   REPEAT     157    187       ANK 5.
FT   REPEAT     201    234       ANK 6.
FT   REPEAT     235    268       ANK 7.
FT   REPEAT     272    301       ANK 8.
FT   REPEAT     303    338       ANK 9.
FT   DOMAIN     364    584       PROTEIN KINASE LIKE.
FT   DOMAIN     585    735       RIBONUCLEASE.
FT   DOMAIN     229    342       2-5A BINDING (P-LOOP) 1.
FT   DOMAIN     263    275       2-5A BINDING (P-LOOP) 2.
FT   ZN_FING    401    416       C6-TYPE (POTENTIAL).
SQ   SEQUENCE   735 AA;  83274 MW;  B6632F4A5B50F711 CRC64;

  Query Match          4.0%;  Score 356;  DB 1;  Length 735;
  Best Local Similarity  32.2%;  Pred. No. 1.5e-11;
  Matches  110;  Conservative   57;  Mismatches  111;  Indels   64;  Gaps   11;

Qy        73 TALTSASKEGHIHIVEELIKSGASLEH RDMGGWTALMWACYKGRTDVVELLLSHGANPS 131
             :: : |.:.:  :.:: : ||      ||| : |   .| |:| :'||||:: 
Db        27 SSLIKAVQKGDVVRVQQLLEKGADANACEDTWGWTPLHNAVQAGRVDIVNLLLSHGADP- 85

Qy       132 VTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLL 191
                                  |   :||'          || : |  :' :: : ::
Db        86 -         -         ---ERRKKNGA- : ------TPFIIAGIQGDVKLLEIIL 112

Qy       192 AMGADVNLQNKSMRALIVAVKQGYTQSVKEILKRNPNVNL---TDKD-------GNTAL 241
             : ||  :     :  : | : |  :::: :  ||.|  |  :       ' |||
Db       113 SDGADVNLKTKTGFTAMMAAREGNAEAIKFLFAKGANVNLRRQTTKDKRRLKQGGATAL 172

Qy       242 MIASREGHIEIVQDLL DAGTYVNIPDRSGDTVLIGAVRGGHV----EIVRALLQKYADI 296
             |:::  :|:::|| :  :  | :  :  :  :  :  :.|   |:  ||:
Db       173 MSAAERGHLEVLRILLNDMKAEVDARDNKGRNALIRTLLNWDCENVEEITSILIQHGADV 232

Qy       297 DIRGQDNKTALYWAVEKGNATMVRDILQ-  CNPDTEICTKDGETPLIKATKMRNIEVVE 353
             :::'::     |  .||: :  :|: :|     | |      :|:| |: |   :  :|:
Db       233 NVRGERGRTPLIAAVERKHTGLVQMLLSREGININDAR--DNEGKTALDIAVDKQLKEIVQ 290

Qy       354 LLLDKGAPVSAVDKKGDTPLHVAIRGRSRRLAELLL --RNP 392
             ||.:|:|     | |  .  :  |     | .''  .|
Db       291 LLLEKGA      DKCDDLVWIARRNHLYHLVKLLLPYVANP 326


RESULT 15
ANR6_HUMAN
ID   ANR6_HUMAN     STANDARD;      PRT;   692 AA.
AC   Q8Y2G4; Q9NU24; Q8IPQ9;
DT   16-OCT 2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat domain protein 6.
```

```
GN   ANKRD6 OR KIAA0957.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain.
RX   MEDLINE=99246063; PubMed=10231032.
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [2]
RP   SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).
RC   TISSUE=Testis;
RA   Poustka A., Klein M., Mewes H.-W., Hassenbuber J., Wiemann S.;
RL   Submitted (SEP 1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 140-692 FROM N.A. (ISOFORM 2).
RA   Tracey A.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9Y2G4-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9Y2G4-2; Sequence=VSP_020706;
CC         Note=No experimental confirmation available.
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB023174; BAA76801.1; -.
DR   EMBL; AL117504; CAB55968.1; -.
DR   EMBL; AL096678; CAB86658.1; -.
DR   HSSP; Q00420; 1ANC.
DR   Genew; HGNC:17280; ANKRD6.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 7.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 9.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Repeat; ANK repeat; Alternative splicing.
FT   REPEAT       8     38       ANK 1.
FT   REPEAT      41     70       ANK 2.
FT   REPEAT      74    104       ANK 3.
FT   REPEAT     107    136       ANK 4.
FT   REPEAT     140    169       ANK 5.
FT   REPEAT     173    202       ANK 6.
FT   REPEAT     206    235       ANK 7.
FT   REPEAT     239    268       ANK 8.
FT   VARSPLIC   264    264       Q -> QVLKPSRPLPLPRRKRSRPLRSSRRALQVERDEVAQSK
FT                               (in isoform 2).
FT                               /FTId=VSP_020706.
FT   CONFLICT    87     87       N -> I (IN REF. 2).
SQ   SEQUENCE   692 AA;  75722 MW;  D3211548FA8B4D5E CRC64;

  Query Match             3.9%;  Score 343.5;  DB 1;  Length 692;
  Best Local Similarity   29.2%;  Pred. No. 1.1;
  Matches  120;  Conservative   67;  Mismatches  177;  Indels   47;  Gaps   11;

Qy         95 ASLEHRDMGGWTALMWACYKGRTDVVEELLSHGANPSVTGLQYSVYPIIWAAIRGHADIV 154
              |:|     |: |  ||:|: :  :: |:  ::|  ::  |:     ::   :

Db          8 AALSER    ---LDVAAYKGQTENVVQLINKGARVAVT--KHGRTPLHLAANKGHLPVV 58

Qy        155 HLLLQNGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKG 214
              :||:   ::  |    | |     :: | :  |:  |  :|::  :  |||   .

Db         59 QILLKAGCDLDVQDDGNQTALHRATVVGNTEIIAALIHEGCALDRQDKDGNTALHEASWH 118

Qy        215 GYTQSVKELLKRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVL 274
              .:||  . :|   ||   :|  :'|| :: :  : :  : ;; ||:  :: : :|.' |

Db        119 GFSQSAKLIVKAGANVLAKNKAGNTALHLACQNSHSQSTRVLLLAGSRADLKNNAGDTCL 178

Qy        275 IGAVRGGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTK 334
                 |   |:  ||  :  :  : |   :||:      :  : :  :'   || :

Db        179 HVAARYNHLSIIKLLLTAPGSVHEKNQAGDTALHVAAALNHKKVAKILLEAGADTTIVNN 238

Qy        335 DGETPLIKATKMRNIEVVELLLDKGAKVSAVDKKGDTP-LHVAIRGRSRRLAELLLRNPK 393
              ':|||       ||  ||    : |    | :'  |:  :   | |  .':

Db        239 AGQTPLETARYHNNPEVA-LLLTKAPQGSV  SAGDTPSSEQAVARKEEAREEFLSASPE 295

Qy        394 - ---DGRLLYRPNKAG----ETPYNIDCSHQKSILTQIFGARHLSPTETD--------- 435
                  |  :||  :       | :   !|:::        :  | : :

Db        296 PRAKDDRRHKSRPKVSAFSDPTPPADQQPGHQKNL------HAHNHPKKRNRHRCSSPPP 349

Qy        436  GDMLGYDLYSSALADILSRPTMQPPICVGLYAQWGSGKSFLLKKLEDEMK 485
               :  |  :  :           ||  :     |     |: |':::::

Db        350 PHEFRAYQLYT--LYRGKDGKVMQAPIN-------GCRCEPLINKLENQLE 391


Search completed: November 21, 2003, 12:45:20
Job time : 25.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

UM protein protein search, using sw model

Run on: November 21, 2003, 11:14:5[illegible] ; Search time 14 Seconds
(without alignments)
342.494 Million cell updates/sec

Title: US 10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLITSQSVINYVEEENIPA[illegible]ENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 818525 seqs, 298052604 residues

Total number of hits satisfying chosen parameters: 818525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score on the left and is derived by analysis of the database.

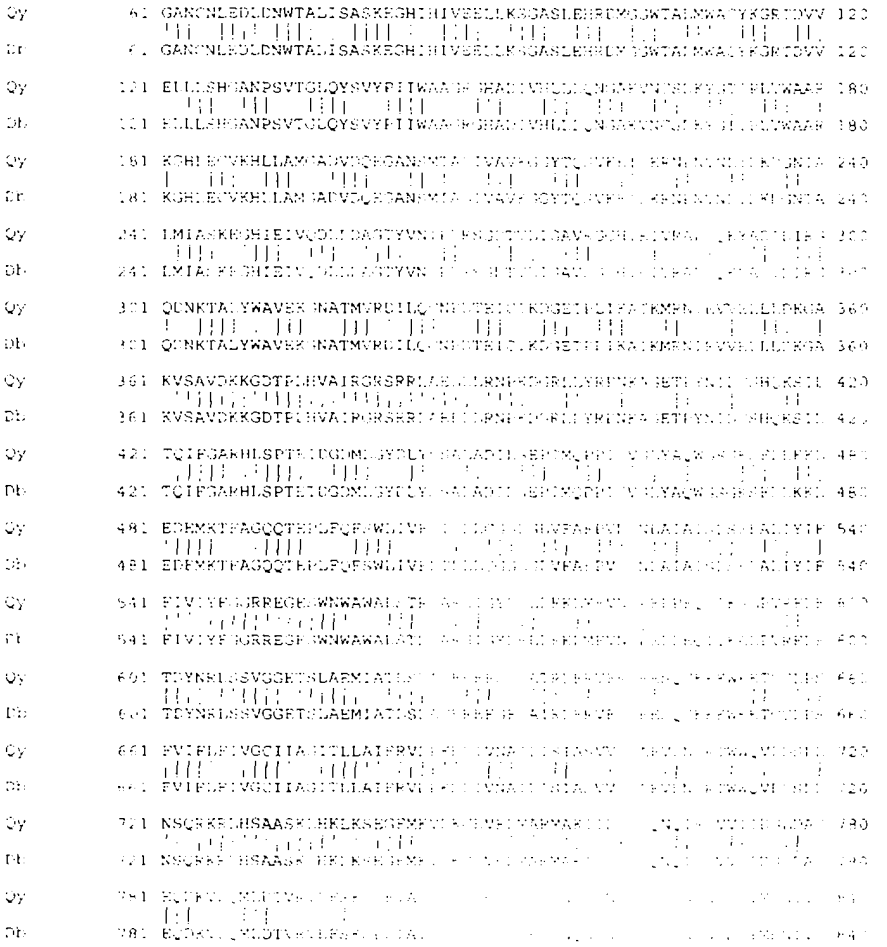
Result No.	Score	Query Match	Length	DB	ID	Description
1	8853	100.0	1715	11	Q9ERD4	Q9erd4 rattus norv
2	8450.5	97.6	1762	11	Q9ERQ9	Q9erq9 rattus norv
3	8212	92.8	1777	4	Q9UBR5	Q9ubr5 homo sapien
4	5073.5	57.3	1031	4	Q9UP44	Q9up44 homo sapien
5	2408	27.2	543	4	Q9H9E8	Q9h9e8 homo sapien
6	2376.5	26.8	1476	5	Q9W2I	Q9w2i drosophila
7	1848.5	20.9	1368	5	Q20107	Q20107 caenorhabdi
8	1097	12.4	239	11	Q8K0I1	Q8k0i1 mus musculu
9	1024	11.6	246	4	Q9NT37	Q9nt37 homo sapien
10	599.5	6.8	129	4	Q9H8E3	Q9h8e3 homo sapien
11	555	6.3	426	16	Q8YT34	Q8yt34 anabaena sp
12	555	6.3	1856	4	Q99437	Q99437 homo sapien
13	551	6.2	1943	11	Q613	Q613 mus musculu
14	545	6.2	1719	4	Q13769	Q13769 homo sapien
15	539	6.1	2622	11	O70511	O70511 rattus norv
16	516.5	5.8	1849	11	Q613	Q613 mus musculu
17	510	5.8	1726	11	Q8VC68	Q8vc68 mus musculu
18	505.5	5.7	1136	6	Q9N180	Q9n180 bos taurus
19	504.5	5.7	1219	11	Q8C6R3	Q8c6r3 mus musculu
20	503.5	5.7	843	11	P97582	P97582 rattus norv
21	500.5	5.7	1486	4	Q8TEF1	Q8tef1 homo sapien
22	499.5	5.6	1159	5	Q9NCP8	Q9ncp8 drosophila
23	493.5	5.6	843	4	Q96I86	Q96i86 homo sapien
24	493.5	5.6	2486	4	O75179	O75179 homo sapien
25	492.5	5.6	4601	5	Q8WRQ7	Q8wrq7 drosophila
26	490.5	5.5	6994	5	Q17343	Q17343 caenorhabdi
27	490.5	5.5	6994	5	Q17490	Q17490 caenorhabdi
28	488.5	5.5	1009	5	Q8SWY2	Q8swy2 drosophila
29	486.5	5.5	1599	11	Q99NH0	Q99nh0 mus musculu
30	486.5	5.5	2542	4	Q8IWZ3	Q8iwz3 homo sapien
31	486.5	5.5	2617	4	Q8IWZ2	Q8iwz2 homo sapien
32	483.5	5.5	1762	11	O88521	O88521 rattus norv
33	479.5	5.4	1188	4	Q9H288	Q9h288 homo sapien
34	479	5.4	2039	5	Q17489	Q17489 caenorhabdi
35	471	5.3	1282	5	Q8T4F8	Q8t4f8 drosophila
36	471	5.3	2119	5	Q9VAU5	Q9vau5 drosophila
37	471	5.3	2119	5	Q8T9G9	Q8t9g9 drosophila
38	467.5	5.3	1549	5	Q9V4B1	Q9v4b1 drosophila
39	464.5	5.2	1549	5	Q24241	Q24241 drosophila
40	459	5.2	743	4	Q8TB46	Q8tb46 homo sapien
41	455	5.1	460	4	Q9H6J9	Q9h6j9 homo sapien
42	454	5.1	1867	5	Q17486	Q17486 caenorhabdi
43	453	5.1	627	4	Q96G77	Q96g77 homo sapien
44	452.5	5.1	1809	5	Q17487	Q17487 caenorhabdi
45	451	5.1	1349	11	Q8VDN6	Q8vdn6 mus musculu

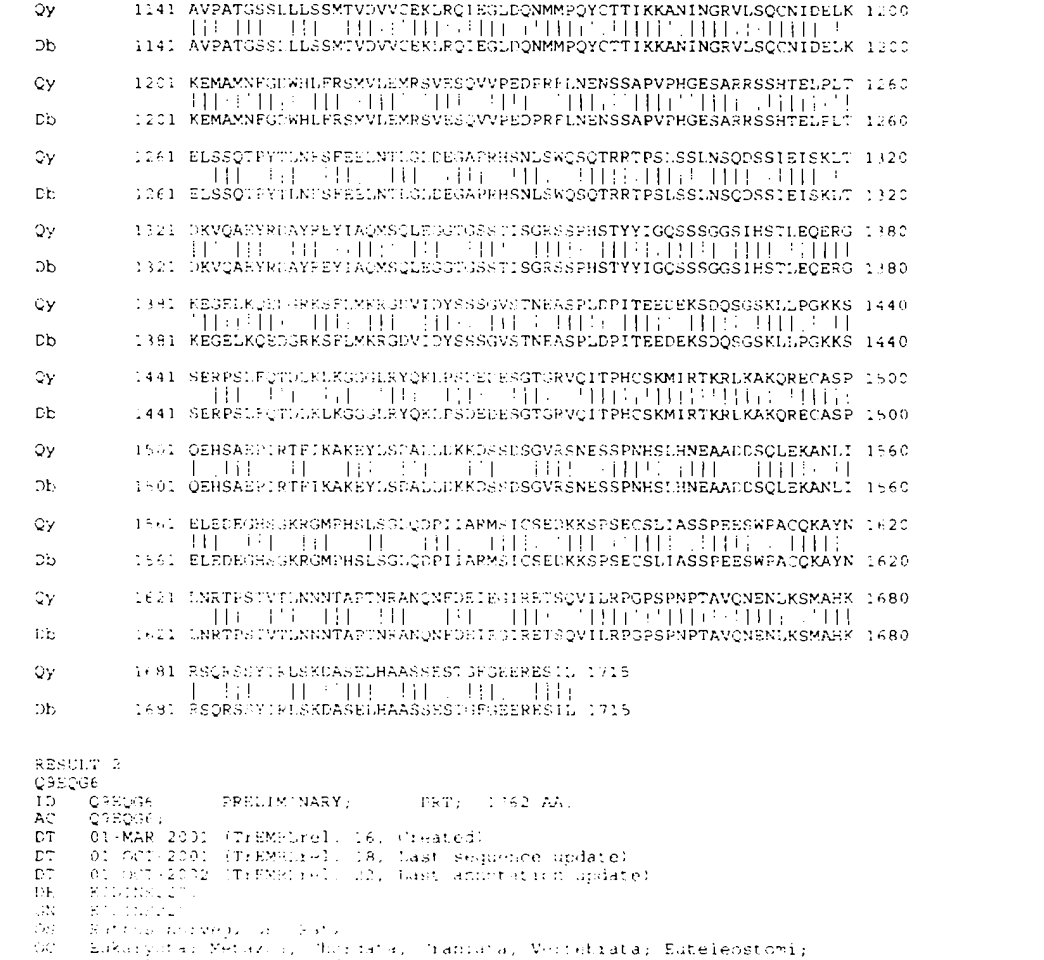
ALIGNMENTS

RESULT 1
Q9ERD4
ID Q9ERD4 PRELIMINARY; PRT; 1715 AA.
AC Q9ERD4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ankyrin repeat rich membrane spanning protein.
GN ARMS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN Sprague-Dawley;
RX MEDLINE=20594524; PubMed=11102474;
RA Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;
RT "An evolutionarily conserved transmembrane protein that is a novel
RT downstream target of neurotrophin and ephrin receptors.";
RL J. Neurosci. 21:176-185(2001).
DR EMBL; AF313464; AAG34167.1; -.
DR HSSP; P80144; 2MYO.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 12.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK REPEAT; 10.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1715 AA; 190542 MW; 505593C4A19A6CDC CRC64;

Query Match 100.0%; Score 8853; DB 11; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSVLITSQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60
Db 1 MSVLITSQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60





```
RESULT 2
Q9EQG6
ID   Q9EQG6      PRELIMINARY;      PRT;  1762 AA.
AC   Q9EQG6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Kidins220.
GN   Kidins220.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20568256; PubMed=10938417;
RA   Iglesias T., Cabrera Poch N., Mitchell M.P., Naven T.J., Rozengurt E.,
RA   Schiavo G.;
RT   "Identification and cloning of Kidins220, a novel neuronal substrate
RT   of protein kinase D.";
RL   J. Biol. Chem. 275:40048-40056(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Schiavo G.;
RL   Submitted (FEB 2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Mitchell M., Schiavo G.;
RL   Submitted (MAY 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF239045; AAG31185.2; -.
DR   HSSP; P80144; 1MY5.
DR   InterPro; IPR002110; ANK.
```

```
DR   Pfam; PF00023; ank; 12.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 10.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1762 AA;  195715 MW;  [illegible]

  Query Match             97.6%;  Score [illegible]  DB 11,  Length 1762;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches 1691;  Conservative    0;  Mismatches   15;  Indels    7;  Gaps    5;

Qy          1 MSVLISQLSVINYVEENIPALKALSEPKEVDEMNKDTQLPLVAAAQQDVEIVELLLKN 60
              ||| |||| |.|| |||  ||  ||| |||  |  | |   |.|
Db          1 MSVLISQMSVINYVEENIPALKALSEPKEVDEMNKDTQLPLMLAAEQGDVEIVKELLKN 60

Qy         61 GANCNLRDLDNWTALISASKEGHIHIVEELLKFGASLEHRDMSWTALMWAAEFGRTDVV 120
              || ' |||| |.|||| |||  ||  ||| |.|  ||' |||  |.  |||
Db         61 GANCNLEDLDNWTALISASKEGHIHIVEILKFGASLEHRDMSWTALMWAAEFGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLIQNGAKVNCSDKYGTTPLVWAAR 180
              |..|||||.|||| |. |||   |||.'||.  ||. ||  ||' |. ||
Db        121 ELLLSHGANPSVTGL-YSVYPIIWAAGFGHADIVHLLIQNGAKVNCSDKYSTTPLVWAAR 179

Qy        181 KGHLECVKHLLAMGADVDCEGANSYTALIVAVEYGYTQAVRELLRNNPNVNIRMKDGNTA 240
              | .|||. |||| ||'|||  | |  ||| |.| |'| || | |||
Db        180 KGHLECVKHLLAMGADVDCEGANSYTALIVAVEYGYTQAVRELLRNNPNVNIRMKDGNTA 239

Qy        241 LMIASKEGHIEIVQDLLDAGTYVNIQNKNDTPLIGAVRSGHVEIVRALLQKYSVDIDIRS 300
              |: |||.| ||||| ''|||  | |  ||  ||. ||  ||  ||  ||
Db        240 LMIASKEGHIEIVQDLLDAGTYVNIQNKNDTPLIGAVRSGHVEIVRALLQKYSVDIDIRS 299

Qy        301 QDNKTALYWAVEKGNATMVRDILQNGITEICKDKFTPLIFATKMPNIEVVELLLFKGA 360
              |||| .|||  |||| ||  ||  ||  | |  |||  |||| |||
Db        300 QDNKTALYWAVEKGNATMVRDILQNGITEICKDKFTPLIFATKMPNIEVVELLLFKGA 359

Qy        361 KVSAVDKKGDTPLHVAIRGRSRRIAELILRNEFTDRDLVYRPNKAGETFYNIDVHGVSIL 420
              |:. ||||| | .|||| |||  ||| ||' ||| ||' | |  ||  |
Db        360 KVSAVDKRGDTPLHVAIRGRSRRIAELILRNEFTDRDLVYRPNKAGETFYNIDVHGVSIL 419

Qy        421 TQIPGARHLSPTETDGDMLDYDLYGSALAPDILSRPTMIPPIGVDLYAQWGFGPFLKKL 480
              ||.  ||||.  |||| ' ||  ||  ||| ||  ||| ||| ||  |
Db        420 TQIPGARHLSPTETDGDMLDYDLYGSALAPDILSRPTMIPPIGVDLYAQWGFGPFLKKL 479

Qy        481 RDEMKTPAGQQTEPLFQFSWLIVPLIIIGGFIGLKVFAMFVDNLIAIALSAVAILVIIF 540
              |||| ''|||   |||  |.|  |.|  ||  ||'  ||  |||
Db        480 RDEMKTPAGQQTEPLFQFSWLIVPLIIIGGFIGLKVFAMFVDNLIAIALSAVAILVIIF 539

Qy        541 FIVIYPLSRRDGRWNNAWAQSLRYHLGMGNSTVENEAFPFDSTHVNLTDQIRERPVIL 600
              ||||'   |||  ||.'  '  '  '  '  '  '  '  '  '  '  '  '  '
Db        540 FIVIYPLSRRDGRWNNAWAQSLRYHLGMGNSTVENEAFPFDSTHVNLTDQIRERPVIL 599

Qy        601 TDYNRLSSVGGETSLAEMIATLSDAVPFGPFKKWFTGTLPSDFVKPKSPTVDTSKIDHE 660
              .|||| .|||  |||  .' |  |  '  '  '  '  '  '  '  '  '  '  '
Db        600 TDYNRLSSVGGETSLAEMIATLSDAVPFGPFKKWFTGTLPSDFVKPKSPTVDTSKIDHE 659

Qy        661 FVIFLPIVGCIIAGITLLAIFRVDFGRLTVNAILSIAAVVLAPVLNQFTWWLVLRSLL 720
              |: ||||||.|||  |  |||  |  |  |  |  |  |  |  |  |  |  |||
Db        660 FVIFLPIVGCIIAGITLLAIFRVDFGRLIVNAILSIAAVVLAPVLNQFTWWLVLRSLL 719

Qy        721 NSQRKRLHSAASKLHKLKSRGPMKVLMSVVELMARMAKTIPPFLQQIRIAVVIDGLDAC 780
              .'|||..|||| ''|||. ||  |||  ||  ||  ||  ||  ||  ||  ||  |
Db        720 NSQRKRLHSAASKLHKLKSRGPMKVLMSVVELMARMAKIIPPFLQQIRIAVVIDGLDAC 779

Qy        781 EQDKVLQMLDTVRVLFSKGPPIAIFASCIHIIIKAINQNLNSVLRDSNINGEIVMENIVH 840
              |  .|| ''|||.|'|||  ||  :  ||  |'|  ||  '|  ||
Db        780 EQDKVLQMLDTVRVLFSKGPPIAIFASCIHIIIKAINQNLNSVLRDSNINGEIVMENIVH 839

Qy        841 LPVFLNSRGLSNARKFLVTSATNGIIVSDTTQIQRDTLRRVSLQSLSFMYKLSNKTALN 900
              || : |||| |||| |'||  |  |  |  |  |  |  |  |  |  |  |  |
Db        840 LPVFLNSRGLSNARKFLVTSATNGIITSDTTQIQRDTLRRVSLQSLSFMYKLSNKTALN 899

Qy        901 RRDTYRRQMQRTITRQMSFDLTKLAVEWHESIRPQTPRRIDIVSVLPIVRANQII 960
```

```
Db        900 RRDTYRRRQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 959

Qy        961 FNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
              | |||| |.||| |||  ||| |||| |||| |  |  |  |  |  |  |||| ||||
Db        960 FNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTIYERISKNIPTTKDVEPLL 1019

Qy        1021 EIDGDIRNFEVFLSSRTPVLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGGLAY 1080
               |||||| ||| '|||. |||| |||| ||||||||||||||||||||'|||||.
Db        1020 EIDGDIRNFEVFLSSRTPVLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGGLAY 1079

Qy        1081 PPLPLHEGPRPPSGYSQPASVCSSASFNGPFPGGVVSPQPHSSYYSGLSGPQHFFYNR- 1139
               ||. |||   |||'|||'|||| '.||||.|||| |||| |'|||| ||||
Db        1080 PPLPLHEGTRPPSGYSQPASVCSSASFNGPFPGGVVSPQPHSSYYSGLSGPQHFFYNRP 1139

Qy        1140 ----                                          ----------AAVPATGSSLLL 1151
                                                            ||||||'||||
Db        1140 FFAPYLYTPRYYPQGSQHLISRSSVKTSLPRDQNNGLPCDSGFNKQRQAAVPATGSSLLL 1199

Qy        1152 SSMTVDVVCEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELKKEMAMNFGDWH 1211
               ||||.||| .|||| ||| |.|||' ||||||||||||||||||||||||||||||||
Db        1200 SSMTVDVVCEKLRQIEGLDQSMMPQYCTTIKKANINGRVLSQCNIDELKKEMAMNFGDWH 1259

Qy        1212 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPLTELSSQTPYTLN 1271
               ||| ||| '|||| ||| |.||| |||| .|||||.||||||||||..|||||
Db        1260 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPLTELSSQTPYTLN 1319

Qy        1272 FSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDKVQAEYRDAY 1331
               ||| |||| |. ||| |.|| |||| | |||||.||||.| ||||||||||||||||||
Db        1320 FSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDKVQAEYRDAY 1379

Qy        1332 REYIAQMSQLEGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERGKEGELKQEDGR 1391
               ||| ||| |||| |||| |||| .|||. .|||||.|||||||||||||||||||||||
Db        1380 REYIAQMSQLEGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERGKEGELKQEDGR 1439

Qy        1392 KSFLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSERPSLFQTDL 1451
               '||| |:| |||' |||' |||'. |||| .|||||.|||||.|..||||'
Db        1440 KSFLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSERPSLFQTDL 1499

Qy        1452 KLKGGSLRYQKLPSDEDESGTGRVQITPHCSKMIRTKRLKAK-QRECASPQEHSAEPIRT 1510
               | ||| ||| |||  |  || . . . | |  || |.. ||||
Db        1500 KLKGGSLRYQKLPSDEDESGTEESDNTP-LLKDDKDKKAEGKAERVCKSP-EHSAEPIRT 1557

Qy        1511 FIKAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADDSQLEKANLIELEDEGHSGK 1570
               ||: |'| '|||' ||| .|||  .||| : |||| '|||| |.|||||.|||
Db        1558 FIKAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADDSQLEKANLIELEDEGHSGK 1617

Qy        1571 RGMPHSLSQLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTPSTVTL 1630
               | |||  ||| |||| |||| ||||  |||| |||||.' ||||||||||
Db        1618 RMPHSLSHLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTPSTVTL 1677

Qy        1631 NSNTAITNPANQNFDEIEGIRETSQVILRPGPSPNPTAVQNENLKSMAHKRSQRSSYTRL 1690
               ||| ||| .|||. ||| |.|| |||| |||||||||||||||||||||||||||||||
Db        1678 NNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQNENLKSMAHKRSQRSSYTRL 1737

Qy        1691 SKDASELHAASSESTGFGEERESIL 1715
               |||| ||| .||| |||||| |||
Db        1738 SKDASELHAASSESTGFGEERESIL 1762


RESULT 3
Q9ULH0
ID   Q9ULH0      PRELIMINARY;      PRT;  1777 AA.
AC   Q9ULH0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein KIAA1250 (Fragment).
GN   KIAA1250.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574461;
RA Nagase T., Ishikawa K., Kikuno R., Hirosawa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033076; BAA86564.2; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 12.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 10.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; ANK repeat; Repeat.
FT NON_TER 1 1
SQ SEQUENCE 1777 AA; 197203 MW; [illegible];

Query Match 92.8%; Score [illegible]; DB 4; Length 1777;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1990; Conservative 93; Mismatches 70; Indels 62; Gaps 4;

Qy 1 MSVLISQSVINYVEEENIPALKAIIEFSDVDSRNEDSQTPLMLAABQGNVEIVKELLKN 60
Db 7 MSVLISQSVINYVEEENIPALKAIIEFSDVDSRNEDSQTPLMLAABQGNVEIVKELIKN 66

Qy 61 GANCNLEDLDNWTALISASKEGHIEIVEILLKNGASLSHRDMDGWTALMWAAYEGKTDVV 120
Db 67 GANCNLEDLDNWTALISASKEGHVEIVEILLKYGVNLEHRDMDGWTALMWAAYEGRTDVV 126

Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAYKGHADIVELLINNGAKVNSGDFYDITPLNWAAR 180
Db 127 ELLISHGANPSVTGLYSVYPIIWAAYKGHADIVELLINNGAKVNSGKYYDITPLNWAAR 186

Qy 181 KGHIEIVKHLLAMGADVDQEGANSMTPIIVAVEGYTQIVKFIIKRNPNVNICDLDNTA 240
Db 187 KGHLEVVKHLLAMGADVDQEGANSMTAIVAVEGYTQVVKFIIKRNPNVNICKCDNTA 246

Qy 241 LMIASKEGHIEIVQDLLDAGIYVNIKDSDGTVLIGAVEGGHVEIVRALLQYKADIIRK 300
Db 247 LMIASKEGHTEIVQDLLDAGIYVNIKDSDGTVLIGAVEGGHVEIVRALLQYKADIIPG 306

Qy 301 QDNKTALYWAVEKQNATVVRDIL [illegible] 360
Db 307 QDNKTALYWAVEKGNATVVRDIL [illegible] 366

Qy 361 KVSAVIFKDQPFLSVAIR [illegible] 420
Db 367 KVSAV [illegible] 426

Qy 421 TQIPGAPHLSPTETDGDMLGYDLV [illegible] 480
Db 427 TQIPGAPHLSPTETDGDMLGYDLV [illegible] 486

Qy 481 EDEMKTPAGQQTEPLFQFSWLIVE [illegible] 540
Db 487 EDEMKTPAGQQIEPLFQFSWLIVE [illegible] 546

Qy 541 PIVIYFIGRRREGESWNWAWVLSTF [illegible] 600
Db 547 PIVIYFIGRRREGESWNWAWVLSTF [illegible] 606

Qy 601 TDYNRLSSVGGETSLAEMIATLSIA [illegible] 660
Db 607 TDYNRLSSVGGETSLAEMIATLSIA [illegible] 666

Qy 661 PVIPLFIVGCIIAGITLLAIFRVCFKHLTVNAVLISIASVVGLAFVLN [illegible] 720
Db 667 PVIPLFIIGCIISGITLLAIFRVCFKHLTVNAVLISIASVVGLAFVLN [illegible] 726

Qy 721 NSQRKRLHSAASKLHKLKSEGFMKVLK [illegible] 780

[illegible]
Db 726 NSQRKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 785

Qy 781 EQDKVLQMLDTVRVLFSKGPFIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Db 786 EQDKVLQMLDTVRVLFSKGPFIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVH 845

Qy 841 LPVPLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900
Db 846 LPVPLNSRGLSNARKFLVTSATNGDVPCSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 905

Qy 901 RRDTYPRPQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960
Db 906 RRDTYRRPQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLPANQIS 965

Qy 961 FNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 966 FNWDRLASWINLTEQWPYRTSWLILYLEETEGIPDQMTLKTIYERISKNIPTTKDVEPLL 1025

Qy 1021 EIDGDIRNFEVFLSSRTPVLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGGLAY 1080
Db 1026 EIDGDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLREIIADVRAAREQISIGGLAY 1085

Qy 1081 PPLPLHEGPPRPPSGYSQPASVCSSASFNGPFPGGVVSPQPHSSYYSGLSGPQHPFYN- 1139
Db 1086 PPLPLHEGPFRAPSGYSQPPSVCSSTSFNGPFAGGVVSPQPHSSYYSGMTGPQHPFYNRP 1145

Qy 1139 ----- - ---- ---------- --- ----------RAAVP 1143
Db 1146 PFAPYLYTPRYYPGGSQHLISRPSVKTSLPRDQNNGLEVIKEDAAEGLSSPTDSSRGSGP 1205

Qy 1144 ATGSSLLLSSMTVDVVCEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELKKEM 1203
Db 1206 APGPVVLLNSLNVDAVCEKLKQIEGLDQSMLPQYCTTIKKANINGRVLAQCNIDELKKEM 1265

Qy 1204 AMNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPLTELS 1263
Db 1266 NMNFGDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTELS 1325

Qy 1264 SQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDKV 1323
Db 1326 SQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDKV 1385

Qy 1324 QAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERGKEG 1383
Db 1386 QAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKGKDS 1445

Qy 1384 ELKQEDGRKNFLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSER 1443
Db 1446 EFKPDDGPFSFLMKRGDVIDYSSSGVSTNCASPLDFITEEDEKSDQSGSKLLPGKKSSER 1505

Qy 1444 PSLFQTDDLKLKGGGLRYQKLPSDEDESGTGRVQITIHCSKMIPTKRLKAKQRECASPQEH 1503
Db 1506 SSLFQTDDLKLKGSGLRYQKLPSDEDESGTEESDNTP-LLKDDKDRKAEGKVERVPKSPEH 1564

Qy 1504 SAEPIRTPIKAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADDSQLEKANLIELE 1563
Db 1565 SAEPIRTPIKAKEYLSDALLDKKDSSDSGVRSSESSPNHSLHNEVADDSQLEKANLIELE 1624

Qy 1564 DEGHSGKRGMPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNR 1623
Db 1625 DDSHSGKRGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNR 1684

Qy 1624 TPSTVTLNNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQNENLKSMAHKRSQ 1683
Db 1685 TPSTVTLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRSQ 1744

Qy 1684 RSSYTRLSKDASELH-AASSESTGFGEERESIL 1715
Db 1745 RSSYTRLSKDPPELHAAASSESTGFGEERESIL 1777

RESULT 4

```
Q9UF42
ID   Q9UF42      PRELIMINARY;      PRT;  1031 AA.
AC   Q9UF42;
DT   01 MAY 2000 (TrEMBLrel. 13, Created)
DT   01-MAY 2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
GN   DKFZP434F0621.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Blum H., Bauersachs S., Mewes H.W., Rieger M., Amann G.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL133620; CAB63746.1; -.
DR   HSSP; P80144; 2MYC.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 10.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   1031 AA;  115311 MW;  [illegible] CRC64;

  Query Match           27.3%;  Score 1478.5;  DB 4;  Length 1031;
  Best Local Similarity   95.2%;  Pred. No.  2;
  Matches  974;  Conservative   23;  Mismatches   13;  Indels   14;  Gaps    2;

Qy          1 MSVLISQSVINYVSEENIPALKALI[illegible]KDVDERNECGQTPLMLAAEQGNVEIVKELIKN 60
              ||'|||||:|||||'|||||'|||:|'|:  |||: |||: ||||: |||  |||  ||:
Db          1 MSVLISQSVINYVSEENIPALKALL[illegible]KDVDERNECGQTPLMLAAEQGNLEIVKELIKN 60

Qy         61 GANCNLEDLDNWTALISASKEGHIHIVEELLK[illegible]GASLEHRDM[illegible]WTALMWAIYKGRTDVV 120
              |||:|||||:|||||'|||||'|||||'|||  |||||'|||||'|||  |||  |||  |||  |||
Db         61 GANCNLEDLDNWTALISASKEGHVHIVEELLK[illegible]GVNLEHRDM[illegible]WTALMWAIYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAIR[illegible]HACIVHILQNGAPVNCSDKY[illegible]TPLVWAAAR 180
              :|:||||':|||||'|||||'|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        121 ELLLSHGANPSVTGLQYSVYPIIWAAIR[illegible]HACIVHILQNGAPVNCSDKY[illegible]TPLVWAAAR 180

Qy        181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKHILKRNPNVNLTDKDGNTA 240
              |||||||'|'|||||'|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKHILKRNPNVNLTDKDGNTA 240

Qy        241 LMIASKEGHIEIVQLLLDAGIYVNIQ[illegible]RQTALHFAAQIGRA 300
              |:  '|'|  |||  ||  |'|  |'|  |'|  |'|  |'|  |'|  |'|  |'|  |'|  |'|  |'|
Db        241 LMIASKEGHTEIVQLLLDAGIYVNIQ[illegible]RQTALHFAAQIGRA 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQ[illegible]TEIGIKDGETPLFIAIRMGNIAVVELLLDKDA 360
              |||  '|:|||||  ||||'|'||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        301 QDNKTALYWAVEKGNATMVRDILQ[illegible]TEIGIKDGETPLFIAIRMGNIAVVELLLDKDA 360

Qy        361 KVSAVDKKGDTPLHVAIRGRSRRLA[illegible]RNPFSGMDLYRPNFAGETPYNIDCSHQPSIL 420
              ||||||'|||||'|:|||||'|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        361 KVSAVDKKGDTPLHIAIRGRSRKLA[illegible]RNPFSGMDLYRPNFAGETPYNIDCSHQPSIL 420

Qy        421 TQIFGARHLSPTETDGDMLGYDLYSAAIADILSEFTMQPPIPVGLYAQWGSGKSFLLKKL 480
              ||||||'|||||'|||||'|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        421 TQIFGARHLSPTETDGDMLGYDLYSAAIADILSEFTMQPPIPVGLYAQWGSGKSFLLKKL 480

Qy        481 EDEMKTFAGQQTEPLFQFSWLIVFLLLLCGGLGLRFAFTVHENLGIAVSLSFLALLYIF 540
              |'||||:|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        481 EDEMKTFAGQQIEPLFQFSWLIVFLLLLCGGLGLRFAFTVHENLGIAVSLSFLALLYIF 540

Qy        541 FIVIYFGGRREGESWNWAWALSTRLARHIGYLELLFKLMFVNFFELPEQTTKALPVRFLF 600
              |'|||||'|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        541 FIVIYFGGRREGESWNWAWVLSTRLARHIGYLELLLKLMFVNFFELPEQTTKALPVRFLF 600

Qy        601 TDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660
```

```
Db        601 TDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660

Qy        661 FVIFLFIVGCIIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720
              |||||||: |||:|'|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKRLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780
              |||||||: |||:|||||'|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        721 NSQRKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy        781 EQDKVLQMLDTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              |||  |||:  |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        781 EQDKVLQMLDTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Qy        841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900
              ||||  |||  |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900

Qy        901 RRDTYRRRQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960
              :||||  |||:|||||  |||:|||  |||  |||  |||  |||  |||  |||  |||  |||  |||:
Db        901 RRDTYRRRQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960

Qy        961 FNWDRLASWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYER-------------IS 1007
              '|||  |||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db        961 FNWDRLASWINLTEQWPYRTSWLILYLEETEGIPDQMTLKTIYERCCGADSCDRDRIGIS 1020

Qy       1008 KNI 1010
              |||
Db       1021 KSV 1023


RESULT 5
Q9H9B4
ID   Q9H9B4      PRELIMINARY;      PRT;   543 AA.
AC   Q9H9B4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ12811.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022873; BAB14285.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   543 AA;  59976 MW;  F5D643D5A20C641D CRC64;

  Query Match           27.2%;  Score 2408;  DB 4;  Length 543;
  Best Local Similarity   86.8%;  Pred. No. 4.3e 155;
  Matches  472;  Conservative   24;  Mismatches   46;  Indels    2;  Gaps    2;

Qy       1174 MMPQYCTTIKKANINGRVLSQCNIDELKKEMAMNFGDWHLFRSMVLEMRSVESQVVPEDP 1232
              |:|||||'||:'||||:|:|||:|| |:|||:| |:|||||||||||||||||||||||||
Db          1 MLPQYCTTIKKANINGRVLAQCNIDELKKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDP 60

Qy       1233 RFLNENSSAPVPHGESARRSSHTELPLTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 1292
              |:|:|:| | | | |:|| ||| |||||||:|'|||||'|||||||||||||||||||
Db         61 RFLSESSSGPAPHGEPARRASHNELPHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 120

Qy       1293 SWQSQTRRTPSLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGTGSSTISG 1352
              ||| |||:|||||||||||||:|||||'|||||'|||||:|||||| ||:|||
Db        121 SWQSQTRRTPSLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISG 180

Qy       1353 RSSPHSTYYIGQSSSGGSIHSTLEQERGKEGELKQEDGRKSFLMKRGDVIDYSSSGVSTN 1412
```

```
Db        181 RSSPHSTYYMGQSSSGGSIHSNDEVEKGKLSEPKPDDGRKSPLMKRGDVIDYRSSGVSTN 240

Qy       1419 EASPLDEITEEDEKSDQSGSKLLM-KKNSERELLPQTDLKLP-SRLRYQPIPSTEDESST 1476
             ||||  ||||| . |||| |  | |  |  ||.  |  || | |  |||
Db        241 DASPLDNITEEDEKSDQSGSKLLMKKPSSERSLLPQTDLKLPKSRLRYQPIPSTEDESST 300

Qy       1477 GRVQITHRCSKMIRTKPLKAKQRE-ASPQRHSADEIRTFIPT-RYLSTA-----TDSS 1532
              ||    |  |  |  |  |  |  | |  |  |  |  |  |  |  |     |||
Db        301 RESDNTP-LLKDDKDRKAESKVEPTLPEHNTAPTEIRKPRYLSDA-----KDLSTGS 350

Qy       1533 VRSNEDPNRSLDNEAADDSQLEPVDIELEDEPSMPLDSEMPLIAPMSIDS 1582
              |||  |||  ||  ||  |  |  |  |  |  |  |  |  |  |||
Db        360 VRSEESPNRSLDNEVADDSHLEPLDIELEDEPSMPLDSEMPLIAPMSIDS 409

Qy       1583 EDKKSPRRCSLIANSPERSWIA-AYNLDNEISTVTLNNLAPNRANLDEIEDLISA 1640
              ||||  |||  |||  |  |  |  |  |  |  |  |  |  |  |  |  ||
Db        410 EDKKSPRRCSLIANSPERNWPACQAYNLDNEISTVTLNNLAPNRANLDEIEDLIST 469

Qy       1641 TSQVILEPGPSPNPTAVQNENLKSMAHRPSQRDSYTPISKDASHLH-AAVEENTGPVEEA 1699
              |||  |||  |  |||  |  |  |  |  |  |  |  |  |  | |  |  |  |
Db        470 TSQVILEPSSSPNPTTIQNENLKSYTHPPSQRDSYTPLSKDPPSLHAAAVEENTGPDEEA 529

Qy       1700 ESIL 1703
              ||||
Db        530 ESIL 533

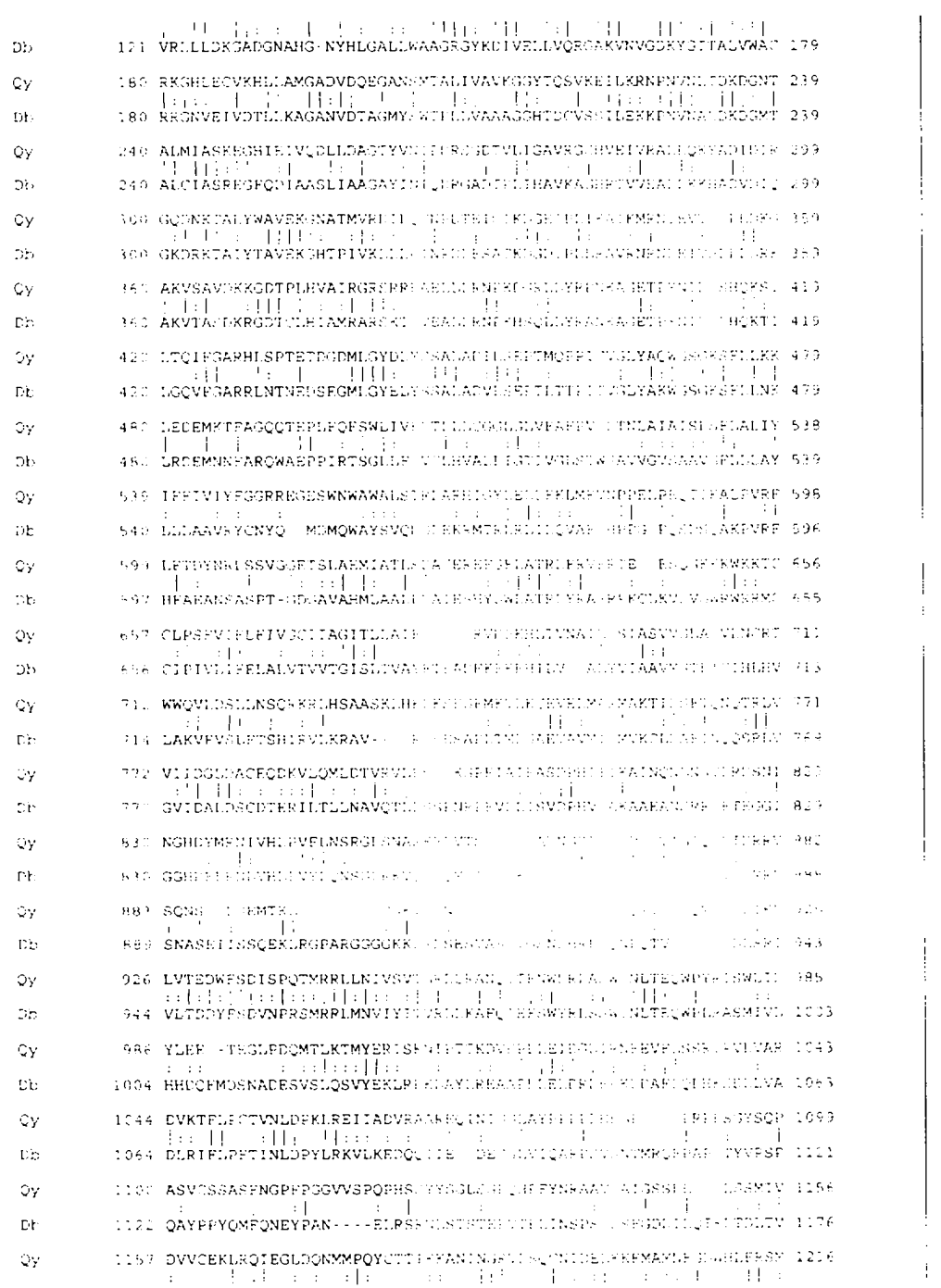

RESULT 6
Q9W2I0
ID   Q9W2I0      PRELIMINARY;      PRT   1476 AA.
AC   Q9W2I0;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG30387 protein
GN   CG30387 OR CG11594 OR CG15669.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN Berkeley;
RX   MEDLINE=20196...; PubMed=1...
RA   Adams M.D., ...
RA   Amanatides P...
RA   George R.A., ...
RA   Sutton G.G., ...
RA   Brandon R.C., ...
RA   Wan K.H., Doyle C., Baxter E.G., ...
RA   Abril J.F., Agbayani A., An H., ...
RA   Ballew R.M., Basu A., Baxendale J., ...
RA   Beeson K.Y., Benos P.V., Berman B.P., ...
RA   Berkowa D., Botchan M.R., Bouck J., ...
RA   Burtis K.C., Busam D.A., Butler H., ...
RA   Cherry J.M., Cawley S., Dahlke C., ...
RA   de Pablos B., Delcher A., Deng Z., ...
RA   Dodson K., Doup L.E., Downes M., ...
RA   Durbin K.C., Evangelista C.C., ...
RA   Fosler C., Gabrielian A.E., Garg N.S., ...
RA   Glodek A., Gong F., Gorrell J.H., ...
RA   Harris N.L., Harvey D., Heiman T.J., ...
RA   Hostin D., Houston K.A., Howland T.J., ...
RA   Jalali M., Kalush F., Karpen G.H., ...
RA   Kimmel B.E., Kodira C.D., Kraft C., ...
RA   Lasko P., Lei Y., Levitsky A.A., ...
RA   Liu X., Mattei B., McIntosh T.C., ...
RA   Merkulov G., Milshina N.V., Moberly C., ...
RA   Mount S.M., Moy M., Murphy B., Murphy L., ...
RA   Nelson D.R., Nelson K.A., Nixon K., ...
RA   Palazzolo M., Pittman G.S., Pan S., ...
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Gara N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paradas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE...; AAF4...; ...
DR   FlyBase; ...; CG30387.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000847; HTH_LysR.
DR   Pfam; PF00023; ank; 11.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 12.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1476 AA;  162594 MW;  19F3376E3F278843 CRC64;

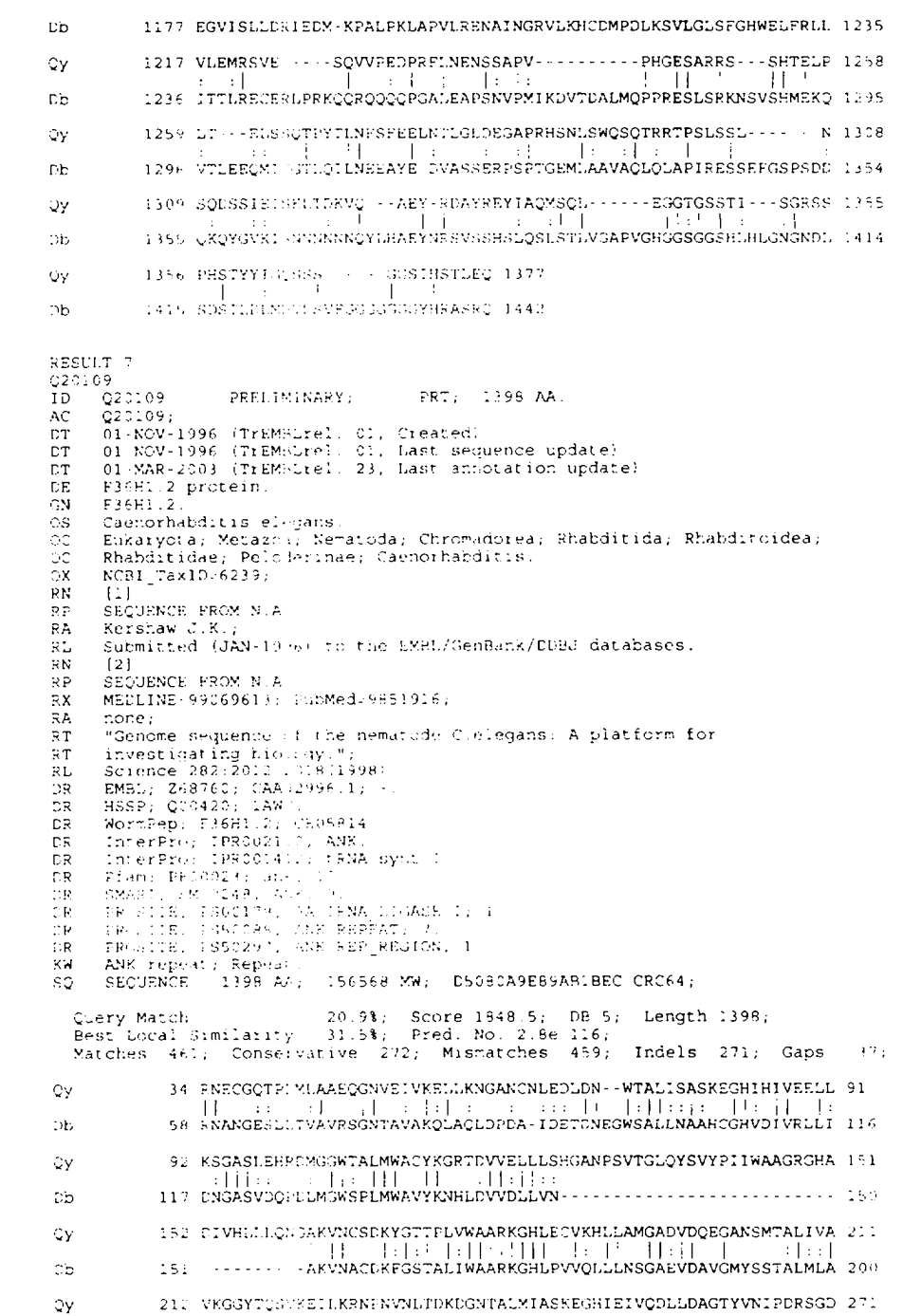
  Query Match             26.8%;  Score 2376.5;  DB 5;  Length 1476;
  Best Local Similarity   37.6%;  Pred. No. 3.5e-152;
  Matches  592;  Conservative  272;  Mismatches  527;  Indels  117;  Gaps   36;

Qy          1 MSVLISQSVINYVEEENIPALKALLE-KCKDVDERNECGQTPLMLAAEQGNVEIVKELLK 59
              |  |  ::::  ::  ::  |:|:|: :  :|:|:|  :  ||: | :|    |:| |
Db          1 MGSLGHRALLQYIDNNDISGLRAILDSRHLTIDDRDENATTVLMVVAGRGLTAFVREFLA 60

Qy         60 NGANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDV 119
               ||:  |  |||  |: |:|: ||: :|: :|  || :|||':||||:|||  ':' |::
Db         61 PGADVQAEDLDNWTALLCASRNGHLDVVQLLLDHGAEVEHRDMGGWTSLMWAAYRGHTEL 120

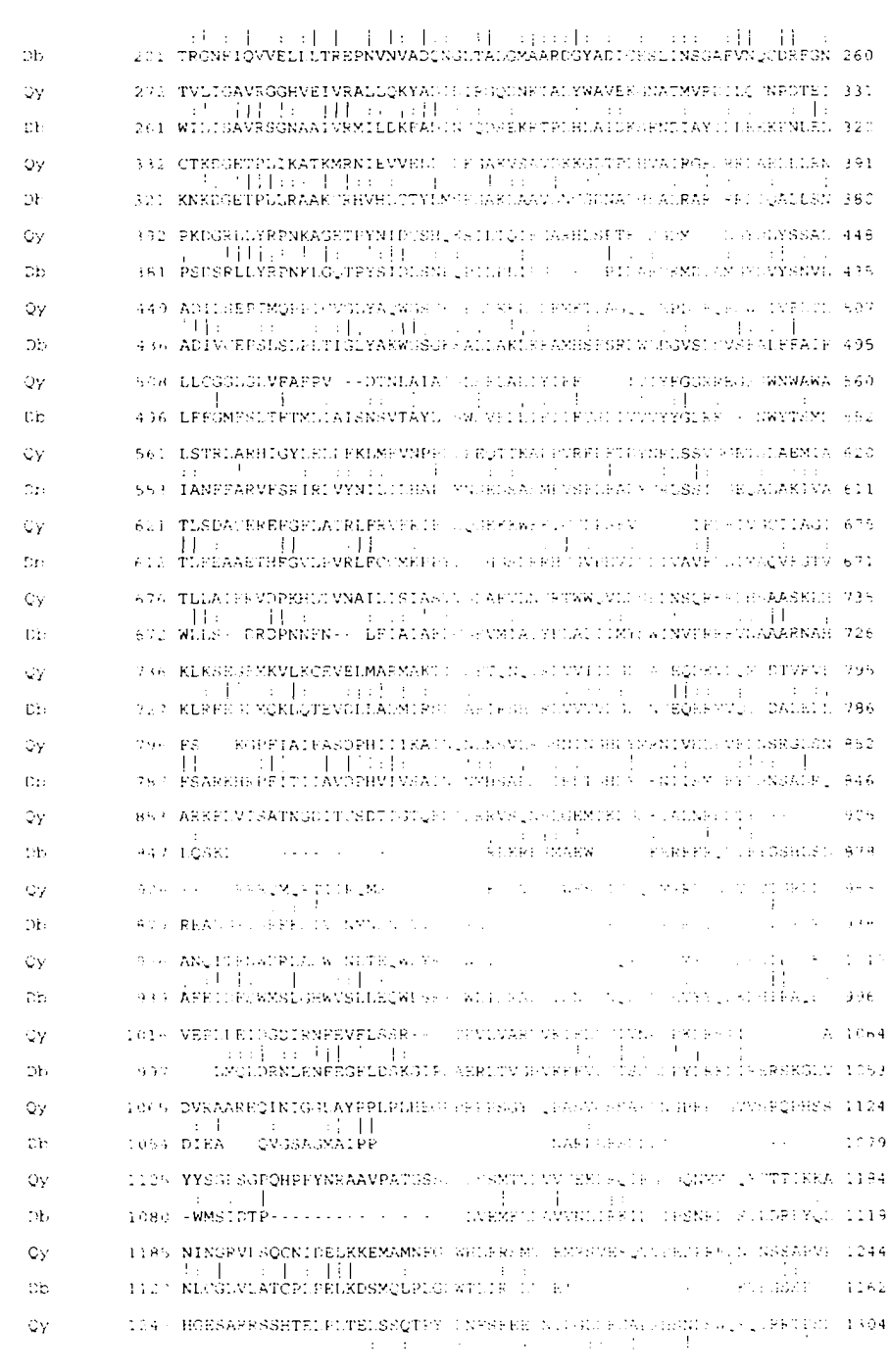
Qy        120 VELLLSHGANPSVTGLQYSVYPRIWAAGRGHADIVRLLLQNGAKVNCSDKYGTTPLVWAA 179
```

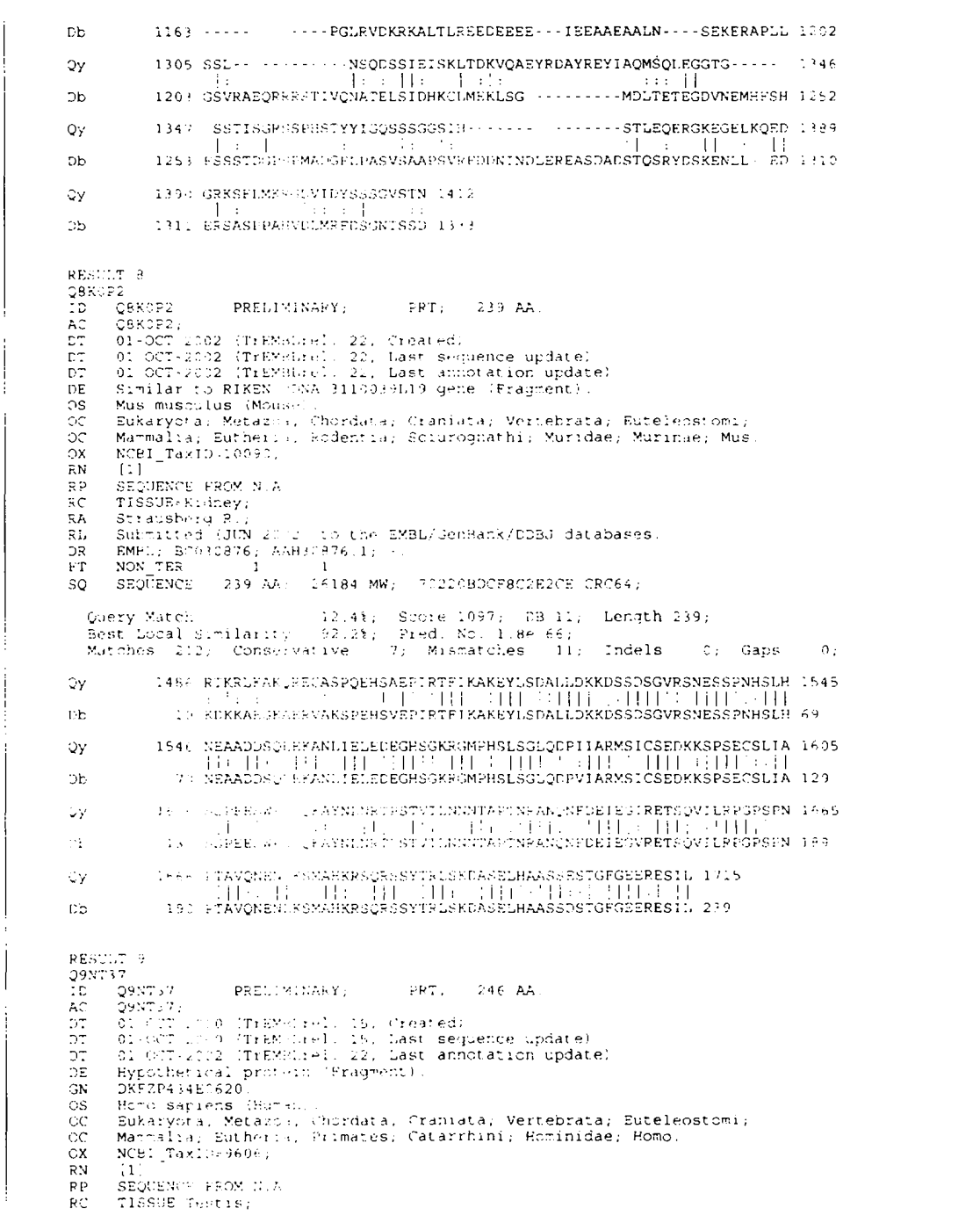




RESULT 7
Q20109
ID Q20109 PRELIMINARY; PRT; 1398 AA.
AC Q20109;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE F36H1.2 protein.
GN F36H1.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (JAN-1[illegible]) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012 - 2018(1998).
DR EMBL; Z68760; CAA92996.1; -.
DR HSSP; Q08420; 1AW7.
DR WormPep; F36H1.2; CE05814.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001412; tRNA synt_1.
DR Pfam; PF00023; ank; 1.
DR SMART; SM00248; ANK; 4.
DR PR; 00118; PS00178; TRNA_LIGASE_1; 1.
DR PR; 00118; PS50088; ANK_REPEAT; 2.
DR PR; 00118; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1398 AA; 156568 MW; D5083A9E89AB1BEC CRC64;

Query Match: 20.9%; Score 1848.5; DB 5; Length 1398;
Best Local Similarity 31.5%; Pred. No. 2.8e-116;
Matches 461; Conservative 272; Mismatches 459; Indels 271; Gaps 47;





RESULT 8
Q8K0P2
ID Q8K0P2 PRELIMINARY; PRT; 239 AA.
AC Q8K0P2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to RIKEN cDNA 3110039L19 gene (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN 2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC010876; AAH10876.1; -.
PT NON_TER 1 1
SQ SEQUENCE 239 AA; 26184 MW; 7D220B0CF8C2E2CE CRC64;

Query Match 12.4%; Score 1097; DB 11; Length 239;
Best Local Similarity 92.2%; Pred. No. 1.8e-66;
Matches 212; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

RESULT 9
Q9NT37
ID Q9NT37 PRELIMINARY; PRT; 246 AA.
AC Q9NT37;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZp434E0620.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;

```
RA   Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Weimann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL137553; CAB70807.1; -.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   246 AA;  27045 MW;  D11B49C628456BF2 CRC64;

  Query Match             11.6%;  Score 1024;  DB 4;  Length 246;
  Best Local Similarity   88.7%;  Pred. No. 1.6e-111;
  Matches  205;  Conservative   15;  Mismatches   19;  Indels    4;  Gaps    3;

Qy       1463 LPSDEGESGTGEVQITPHCSKMIRELKAKQSEDASPQFRSAPDIRTPFTAFEYISDAL 1522
             |  |.     .:  :  |                     |    ' |   ||   .:  ':|
Db         12 LKDDKDRKAESKVEAVP-  --                KSP EHSAAPIRTPFSAFAYISDAL 52

Qy       1523 LDKKDSSDSGVRSNESSPNHSLHNFAADDNQLEKANLIELSDEGHSGKKAMIKSGLQD 1582
             ||||:''|||  :|||:'|||  '||:'|'|  :||  :  |  ||:   :|  ||:
Db         53 LDKKDSSDSGVRSSESSFNHSLHNFVADDNHLEFANLIELEDDSHSGKRTIHSLSGLQD 112

Qy       1583 PIIAEMSICSEDKKSPSECSLIASSPEESWFACQKAYNLNRTPSTVILNNNTAPTNRANQ 1642
             : |||:.'|||| ' ||:|| | '||:  :|:|  ||| : | |||  :.  ||'
Db        113 PIIAEMSICSEDKKSPSECSLIASSPEENWFACQKAYNLNRTPSTVILNNNSAPANRANQ 172

Qy       1643 NFDEIEGIRETSQVILRPGPSPNPTAVQNENLKSMAHKRSQRSSYTRLSKDASELH-AAS 1701
             |  |.|| ; |||||:''  |||   :'|   |||  ||'' ||.' ||;   .|  |||
Db        173 NFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRSQRSSYTRLSKDIPELHAAAS 232

Qy       1702 SESTGFGEERESIL 1715
              |||,  ''|||  :
Db        233 SESTGFGEERESIL 246


RESULT 10
Q9H889
ID   Q9H889      PRELIMINARY;      PRT;   129 AA.
AC   Q9H889;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ13864
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RA   Isogai T., Ota T., Hayashi K., Suzuyama T., Isuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Isogai T., Hara M.,
RA   Tanase T., Nomura Y., Tsuiya S., Ishii S., Sugiyama T., Saito K.,
RA   Arita M., Nakamura T., Ishii S., Kawai J., Carninci P., Hayashizaki Y.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Yamashita R., Ishida A.,
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK023926; BAB14728.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   129 AA;  14244 MW;  4BCF4168137485B4 CRC64;

  Query Match              6.8%;  Score 599.5;  DB 4;  Length 129;
  Best Local Similarity   90.7%;  Pred. No. 4.4e-58;
  Matches  117;  Conservative    4;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1598 MSICSEDKKSPSECSLIASSPEESWFACQKAYNLNRTPSTVILNNNTAPTNRANQNFDEM 1657
             |.||||:   |||:|'|||.:'||:   |||  :|| :'||:  ||': |    ||:|:||:
Db          1 MSICSEDKKSPSECSLIASSPEENWFACQKAYNLNRTPSTVILNNNSAPANRANQNFDEM 60

Qy       1658 EGIRETSQVILRPGPSPNPTAVQNENLKSMAHKRSQRSSYTRLSKDASELH-AASSESTG 1716
             :'|||:' ||||  |||||:||   |||:'||:  |||  |':  |':  |||' ||
Db         61 EGIRETSQVILRPSSSPNPTTIQNFNLKSMTHKRSQRSSYTRISKDIPELHAAASSESTG 120

Qy       1717 FGEERESIL 1725
              ||||||''
Db        121 FGEERESIL 129


RESULT 11
Q8YTG9
ID   Q8YTG9      PRELIMINARY;      PRT;   426 AA.
AC   Q8YTG9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein All2748.
GN   All2748.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; AP003590; BAB74447.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   426 AA;  43800 MW;  CF0B9B5D43ACAA03 CRC64;

  Query Match              6.8%;  Score 555;  DB 16;  Length 426;
  Best Local Similarity   47.2%;  Pred. No. 4.1e-29;
  Matches  149;  Conservative   67;  Mismatches  177;  Indels    8;  Gaps    6;

Qy          9 VINYVREENIPALKALLEKGKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNL-E 67
              ::  : .|   |||     |: : :  || ,:  |  |:|: ||  .'  ||
Db          8 LLKVAKSGDIKGLGALLAAGVGVDICDRDGTTALMFAANLGYTEIVRSLLDGGANVNLAR 67

Qy         68 DLDNWTALISASKEGHIRIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVVELLLSHG 127
                 ||: :|:   :  :: |: :|::   : | '||| |  |:  :|  : |: |
Db         68 KRYGLTALMLAASANQVDIVQLLISRGAAVNATNEDGSTALMAAALKGNVEVARVLLAAG 127

Qy        128 ANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKGHLECV 187
              :: :::   :  | |  | | :  |:' :.|  |:.,: | | :: ||  ', : |
Db        128 ADVNITDRDDT-ALKLAVKRGQAAVVQLILPSGADANCEDEEGETLLMLAADSGHGDVV 186

Qy        188 EELLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTALMIASKE 247
              : :|| |  .::  :  |:' || .      :| |  :|| |:|' :|| :|: |
Db        187 QVLLAAGVEVNEQNQDGGTALLAAVAAGNGAIAKILLDRGADVNHQDQDGESALHLATVE 246

Qy        248 GRIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQD-NKTA 306
              |:::|:  : |   | :: |||  : |  ||:|| :| ||   :: | :  :|
Db        247 GYVDVVQVLLNQGANTQIKNKLGDTPLDVAALQGHDQIVETLL-KYG-ANVHGDNLGETP 304

Qy        307 LYWAVEKGLATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGAKVSAVD 366
              |    :|:  || :|   : |  ||:| :|||:  :  |::||| ||| |:  |
Db        305 LTLAASQGHIATVRILLDYGANANIPASDGKTALIKATERNHPGVIQLLLAKGANVNYQD 364

Qy        367 KKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGET 407
              ! ! |    |  :: ::||   |  :    :| | |
Db        365 SVGATALIWAASGGYNKVVQILLEGGADTNL---KNRGGYT 402


RESULT 12
Q99407
ID   Q99407      PRELIMINARY;      PRT;  1856 AA.
AC   Q99407;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
```

```
DT   01 MAY 1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ankyrin
GN   ANK1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97382244; PubMed=9215714;
RA   Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;
RT   "Structure and organization of the human ankyrin-1 gene and
RT   complexity of pre-mRNA processing.";
RL   J. Biol. Chem. 272:19220-19228(1997).
DR   EMBL; U50133; AAB47805.1; -.
DR   EMBL; U50092; AAB47805.1; JOINED.
DR   EMBL; U50093; AAB47805.1; JOINED.
DR   EMBL; U50094; AAB47805.1; JOINED.
DR   EMBL; U50095; AAB47805.1; JOINED.
DR   EMBL; U50096; AAB47805.1; JOINED.
DR   EMBL; U50097; AAB47805.1; JOINED.
DR   EMBL; U50098; AAB47805.1; JOINED.
DR   EMBL; U50099; AAB47805.1; JOINED.
DR   EMBL; U50100; AAB47805.1; JOINED.
DR   EMBL; U50101; AAB47805.1; JOINED.
DR   EMBL; U50102; AAB47805.1; JOINED.
DR   EMBL; U50103; AAB47805.1; JOINED.
DR   EMBL; U50104; AAB47805.1; JOINED.
DR   EMBL; U50105; AAB47805.1; JOINED.
DR   EMBL; U50106; AAB47805.1; JOINED.
DR   EMBL; U50107; AAB47805.1; JOINED.
DR   EMBL; U50108; AAB47805.1; JOINED.
DR   EMBL; U50109; AAB47805.1; JOINED.
DR   EMBL; U50110; AAB47805.1; JOINED.
DR   EMBL; U50111; AAB47805.1; JOINED.
DR   EMBL; U50112; AAB47805.1; JOINED.
DR   EMBL; U50113; AAB47805.1; JOINED.
DR   EMBL; U50114; AAB47805.1; JOINED.
DR   EMBL; U50115; AAB47805.1; JOINED.
DR   EMBL; U50116; AAB47805.1; JOINED.
DR   EMBL; U50117; AAB47805.1; JOINED.
DR   EMBL; U50118; AAB47805.1; JOINED.
DR   EMBL; U50119; AAB47805.1; JOINED.
DR   EMBL; U50120; AAB47805.1; JOINED.
DR   EMBL; U50121; AAB47805.1; JOINED.
DR   EMBL; U50122; AAB47805.1; JOINED.
DR   EMBL; U50123; AAB47805.1; JOINED.
DR   EMBL; U50124; AAB47805.1; JOINED.
DR   EMBL; U50125; AAB47805.1; JOINED.
DR   EMBL; U50126; AAB47805.1; JOINED.
DR   EMBL; U50127; AAB47805.1; JOINED.
DR   EMBL; U50128; AAB47805.1; JOINED.
DR   EMBL; U50129; AAB47805.1; JOINED.
DR   EMBL; U50130; AAB47805.1; JOINED.
DR   EMBL; U50131; AAB47805.1; JOINED.
DR   EMBL; U50132; AAB47805.1; JOINED.
DR   HSSP; Q80422; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR003906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1856 AA;  203466 MW;  AA4B3A7940174401F 7B 94,
```

```
  Query Match          6.3%;  Score 555;  DB 4;  Length 1856;
  Best Local Similarity  20.6%;  Pred. No. 5.6e-28;
  Matches  395;  Conservative  269;  Mismatches  644;  Indels  612;  Gaps   74;

Qy        21 LKALLEKGKDVDERNEDGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWTALISASK 80
             :: |:    :|: :::   ||  :||:: ::|:: |||:|| | ||:   | :  | |:
Db        61 VRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQ 120

Qy        81 EGH----  - - - ------ - ---------- ---------------------- 82
             :
Db       121 QGHENVVA-LLNYGTK-GKVRLPALHIAARNDDTRTAAVLLQNDPNPDVLSKTGFTPLHIA 180

Qy        84     - IR:VRELLKSGASIEHRDMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYS 138
                 ::   :  || :  | |   |  :    :: ||  -  :  :
Db       181 AHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIE-TKTKDE 239

Qy       139 VYPIIWAA-MGHADIVHLLLQNGAKVNAFSKYGTTPLYWAARKGHLECVKHLLAMGADVD 198
             : :  |.  |:    :  :|  .  :  | :|:  ||:  .|:||.  |   :  :|
Db       240 LTPLHCAARNGHVRISELLLDHGADIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYDAEID 299

Qy       199 -------    ---- --- -    -QEGA------ NSMTALIVAVKGGYTQSVKEILK 235
                                      :||          | | :  |  : : : : :
Db       300 DITLDHLTPLHVAAHCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRVMELLLK 359

Qy       236 RNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEI 295
             :  :::   :  | | :  :|:  ||:  : ::|   |    : :   :| : | || ':
Db       360 TGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEV 419

Qy       296 VRALLQKYAVDIDRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATK 355
             : |   |  : : : :::|  : |   |:  ' : :|: ' :  : | |  |  | |
Db       420 AKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAR 479

Qy       356 MRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRPLAELLLRNPKDGRLDYRPNKAG 415
               ::|    | :  |  : :   | |  || :    :|: ||   :       .  ||
Db       480 EGHVETVLALLEKEASQACMTKKGFTPLHVAAKYGKVRVAELLLE--RDA    HPNAAG 533

Qy       416 N---TPYNIDCSHQ-----KSIL- - ------------------TQIFGAR------  457
             :   || ::          | :                        |: ||
Db       534 KNGLTPLHVAVHHNNLDIVKLLLPRGGSPHSPAWNGYTPLHIAAKQNQVEVARSLLQYGG 593

Qy       458 ---------- --HLSPTETDGDMLGYDLYSSALADILSEPTMQP------- --PIC-  502
                        '|:    : :   :  |  : : :  : |              |:
Db       594 SANAESVQ-PVTPLHLAAQEGHAEMVALLLSKQANGNLGNKSGLTPLHLVAQEGHVPVADV 653

Qy       503 -- --       ------VGLYAQWGSGK--SFLLKKLED-EMKTFAGQQTEPLFQFSWL 551
                          : : : : |  |||:  |   :|  |    :| | :
Db       654 LIKHGVMV-ATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNAKTKLGY--SPLHQAAQQ 711

Qy       552  -IVFIT  :  GGAGLVFAFFVD      TNLAIAISLSFLALIYIFFIVIYHGGRREGE 603
                  ,:   |  :     |  '|     :::  :  :|
Db       712 DHIDIVE  KNG     ASPNGVSSDFTTPLAIAPFLGYISVTDVLKVV          756

Qy       604 SWNWAWAA  LPLAPHIGVIELLFKIVFVNIFFLPFQPTKALPVPFLFTDYNPLS SVGGE 663
                 :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db       757 TDETSFVLVSDKHR-MSFPETVDELDVSEDEGTAHITIMGEELISFKAERRDSRDVDEE 815

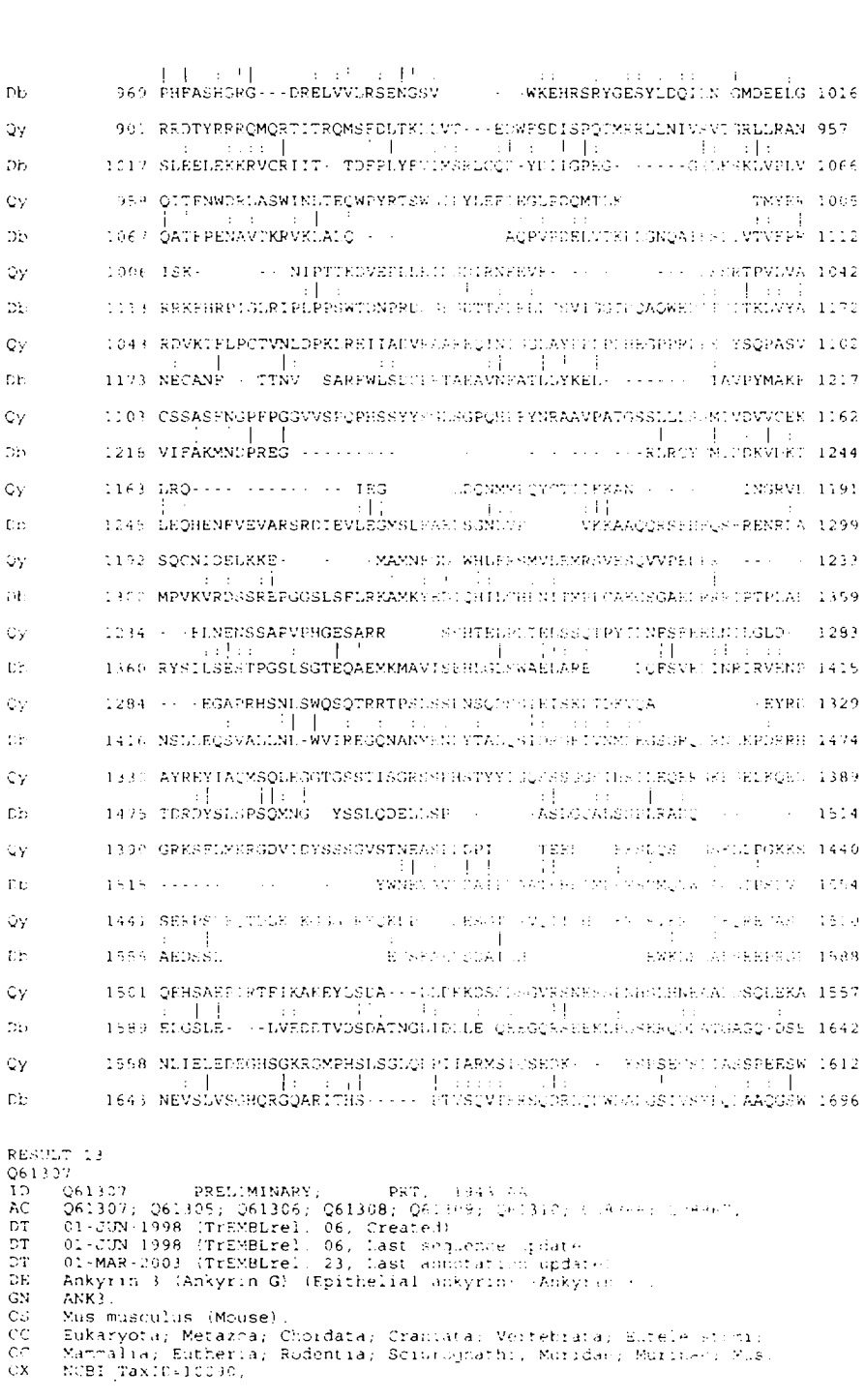
Qy       613 TSLAEMIATLSDACEREPGPLATRLFR---------VFRTEE-SQGKKKWKKTCCLFSFV 662
               | :          :  : |          ' |:|:  |  |:: :  :||
Db       816 KELLDFVIKLDQVVE------SPAIPRIPCAMPETVVIRSEEQEQASKEYDEDSLIPS-- 867

Qy       663 IFLPIVGFIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLLNS 722
                                 |  | :  :  :|  |  |:::      :::   |
Db       868 -------        -------SPATETSDNI-SPVASPVHTGFLVSF-----MVDARGGS 900

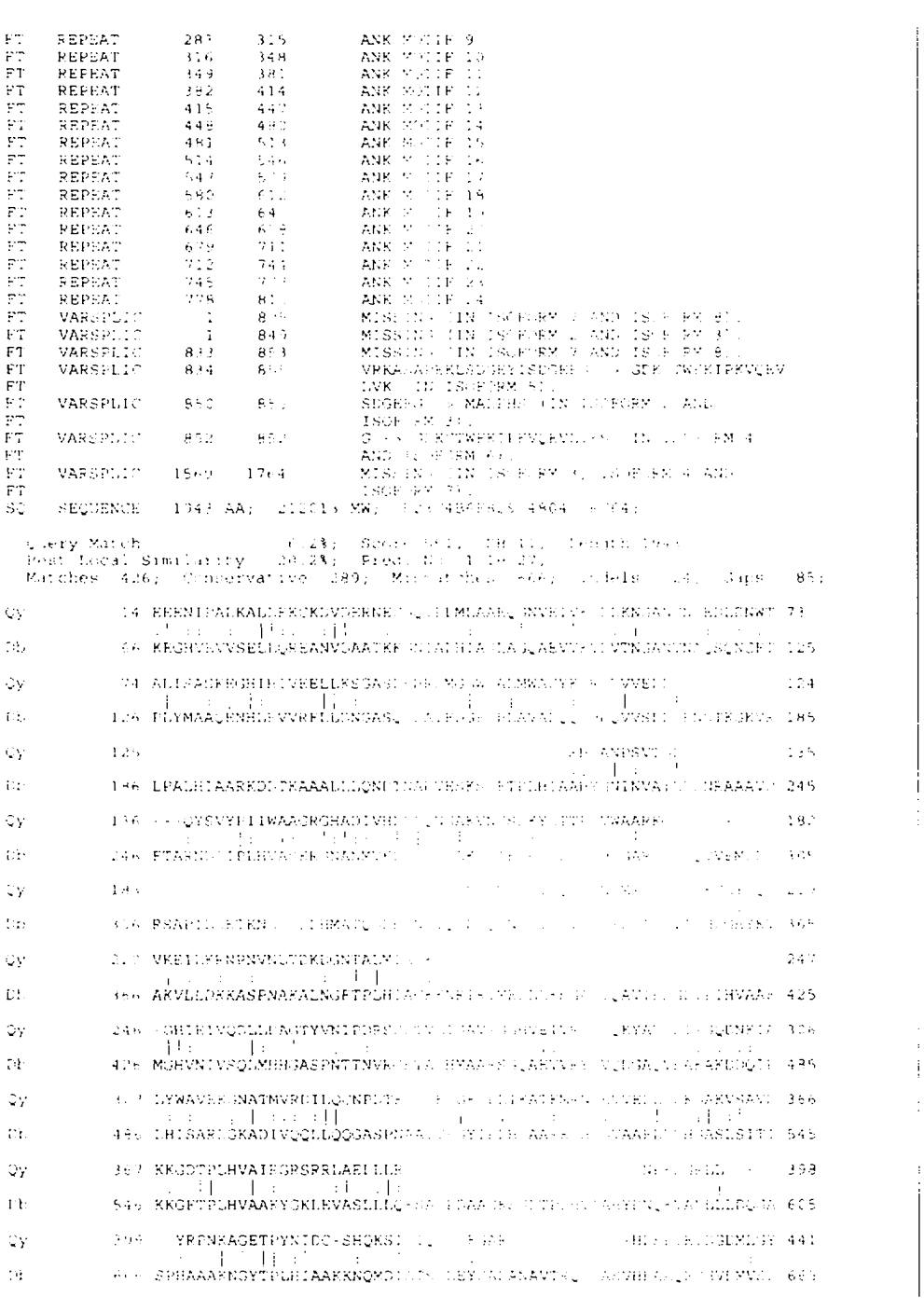
Qy       723 QRKRLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDACEQ 782
                 ::                |  |:          :|  :  |:   |||
Db       901 MPGSPHN  -------- -GLRVVIP---------PRTCAAPTRITCRLV---------- 929

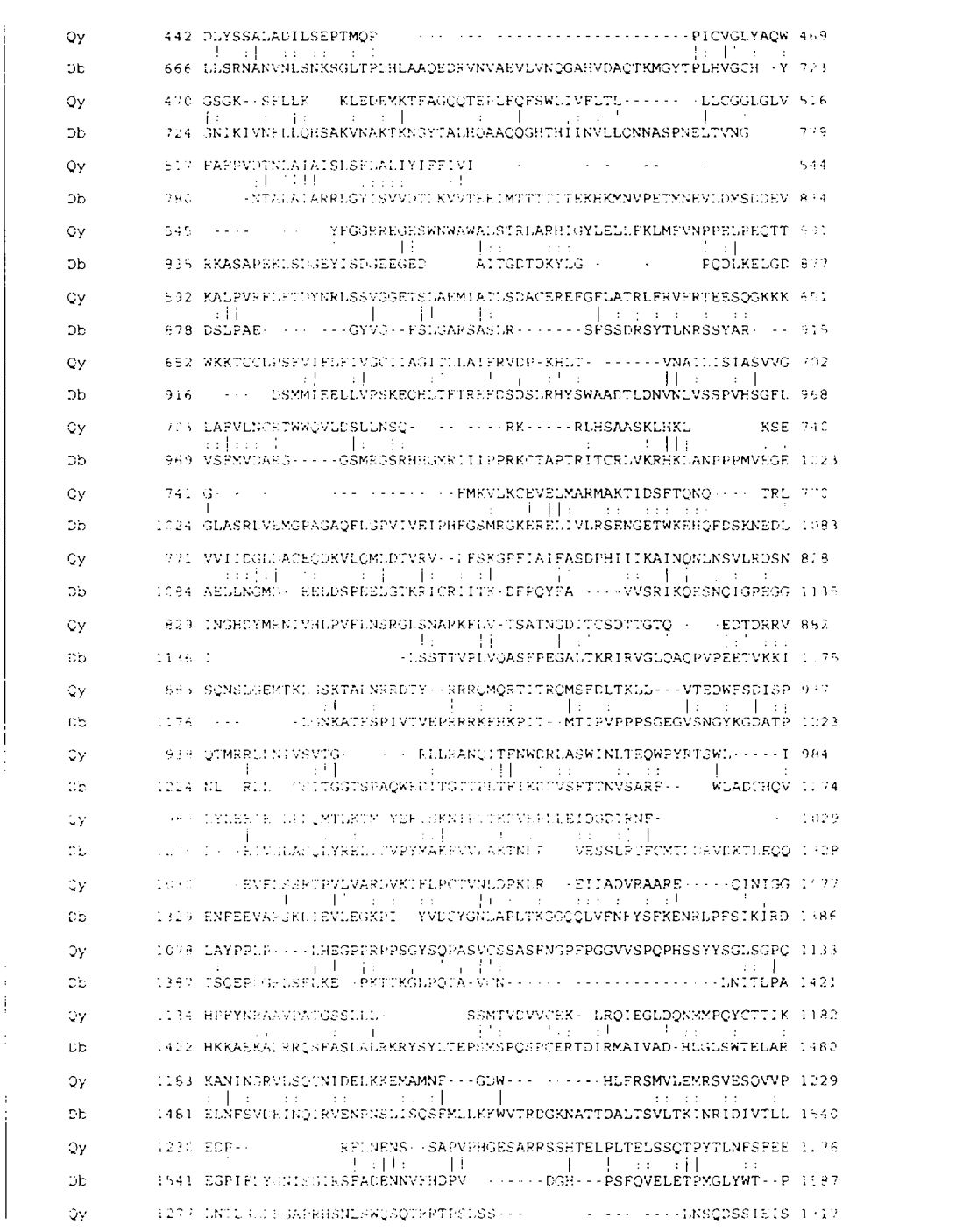
Qy       783 DKVLQMLPTVRVLFSKGPPIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRN-IVHL 841
               |  :  |  |                  |  : | |  :      |  ::    || :
Db       930   KPQKLSTPPPL---   ------------AEEEGLASRIIALGPTGAQFLSPVIVEI 968

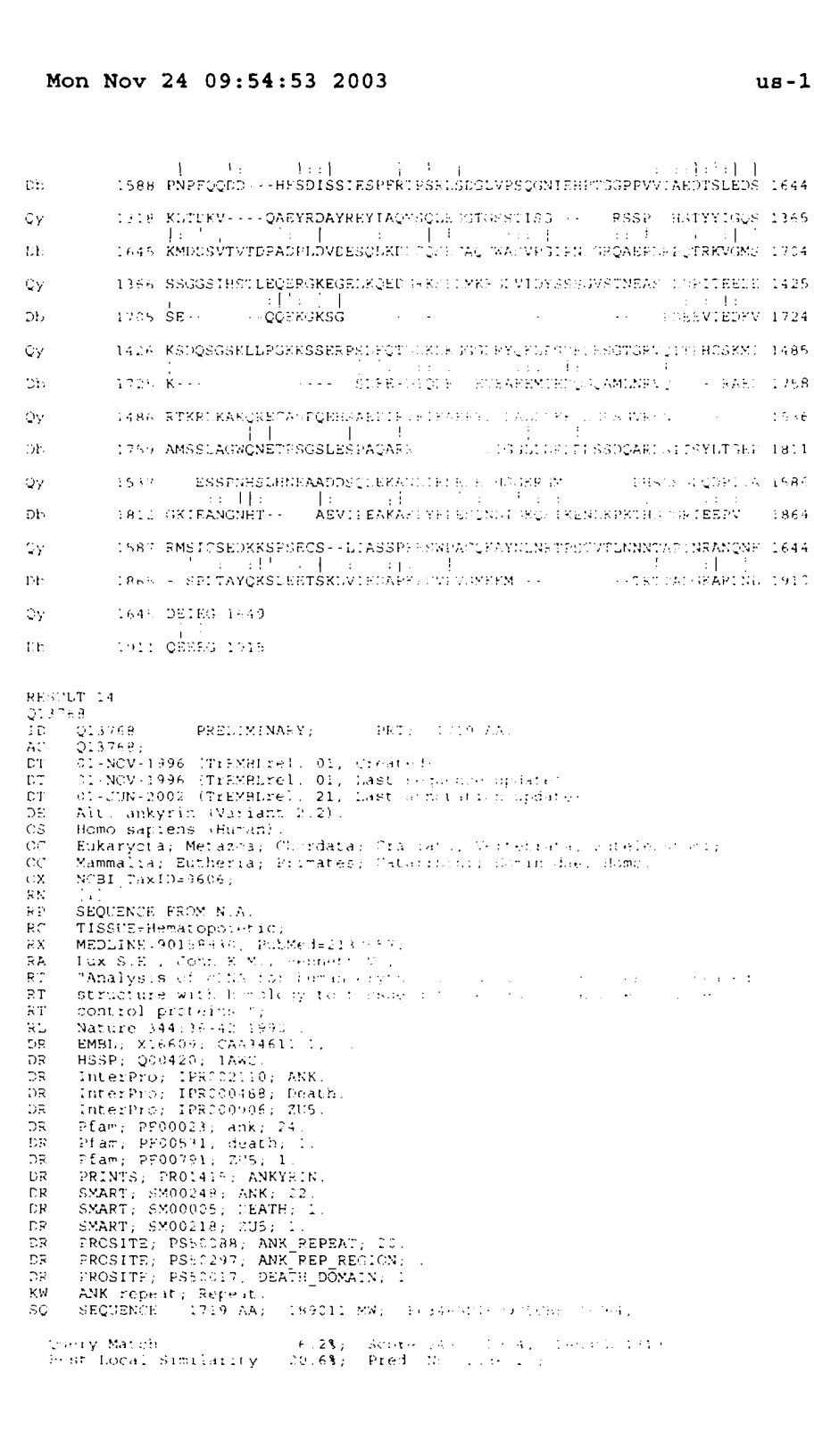
Qy       842 PVFL-NSP DLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900
```



```
RESULT 13
Q61307
ID   Q61307      PRELIMINARY;      PRT,  1883 AA.
AC   Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; Q61311;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ankyrin 3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin-3).
GN   ANK3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=95340633; PubMed=7615634;
RA   Peters L.L., John K.M., Lu F.M., Eicher E.M., Higgins A., Yialamas M.,
RA   Turtzo L.C., Otsuka A.J., Lux S.E.;
RT   "Ank3 (epithelial ankyrin), a widely distributed new member of the
RT   ankyrin gene family and the major ankyrin in kidney, is expressed in
RT   alternatively spliced forms, including forms that lack the repeat
RT   domain.";
RL   J. Cell Biol. 130:313-330(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RC   STRAIN=C3H/HeJ; TISSUE=Bone Marrow;
RX   MEDLINE=97213781; PubMed=9060472;
RA   Hoock T.C., Peters L.L., Lux S.E.;
RT   "Isoforms of ankyrin-3 that lack the NH2-terminal repeats associate
RT   with mouse macrophage lysosomes.";
RL   J. Cell Biol. 136:1059-1070(1997).
CC   -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
CC       SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
CC       BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
CC       AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
CC       DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
CC       THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC       THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC       ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1 (SHOWN HERE), 2, 3,
CC       4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC       EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC       LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC       SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC   -!- SIMILARITY: CONTAINS 24 ANK REPEATS.
DR   EMBL; L40631; AAB01603.1; -.
DR   EMBL; L40632; AAB01605.1; -.
DR   EMBL; U89275; AAB58391.1; -.
DR   EMBL; L40631; AAB01602.1; -.
DR   EMBL; L40632; AAB01604.1; -.
DR   EMBL; L40632; AAB01606.1; -.
DR   EMBL; L40632; AAB01607.1; -.
DR   EMBL; U89274; AAB58390.1; -.
DR   HSSP; P55273; 1BI8.
DR   MGD; MGI:88026; Ank3.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001452; Death.
DR   InterPro; IPR000210; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; Death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Multigene family.
FT   DOMAIN        1    816       ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT   DOMAIN       23    815       24 X ANK MOTIF REPEATS.
FT   DOMAIN      856   1437       SPECTRIN BINDING DOMAIN.
FT   DOMAIN     1438   1943       REGULATORY DOMAIN
FT                                (REGULATES BINDING OF ANKYRIN TO SPECTRIN
FT                                AND THE BAND 3 PROTEIN).
FT   REPEAT       23     55       ANK MOTIF 1.
FT   REPEAT       56     88       ANK MOTIF 2.
FT   REPEAT       89    121       ANK MOTIF 3.
FT   REPEAT      122    154       ANK MOTIF 4.
FT   REPEAT      155    183       ANK MOTIF 5.
FT   REPEAT      184    216       ANK MOTIF 6.
FT   REPEAT      217    249       ANK MOTIF 7.
FT   REPEAT      250    282       ANK MOTIF 8.
```



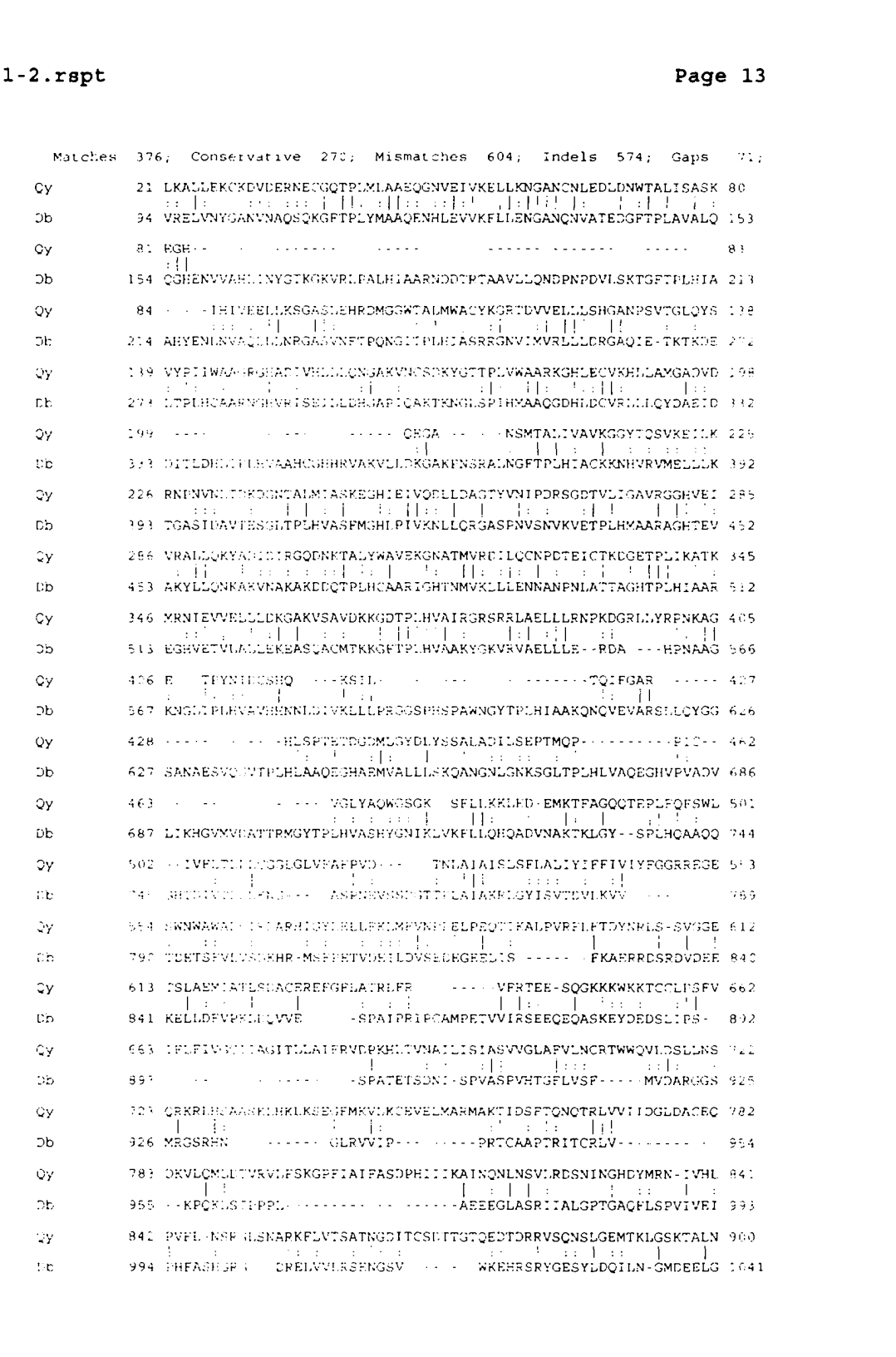


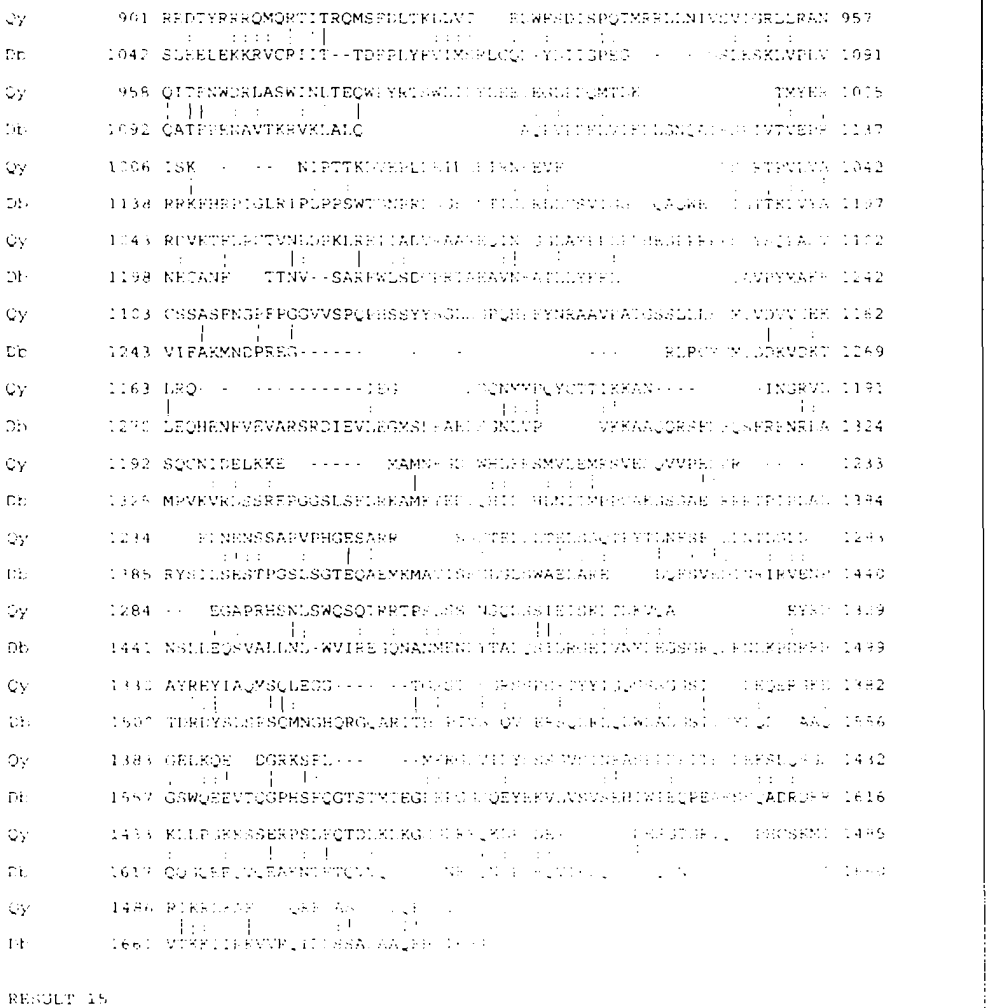


RESULT 14
Q13768
ID Q13768 PRELIMINARY; PRT; 1719 AA.
AC Q13768;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Alt. Ankyrin (Variant 2.2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hematopoietic;
RX MEDLINE=96158446; PubMed=8617211;
RA Lux S.E., Corn K.M., Menguy T.,
RT "Analysis of cDNAs for human erythroid ankyrin.";
RT Structure with homology to yeast cell cycle proteins.";
RT control proteins.";
RL Nature 344:36-42 1990.
DR EMBL; X16609; CAA34611.1.
DR HSSP; Q00420; 1AWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1719 AA; 189011 MW; [illegible];

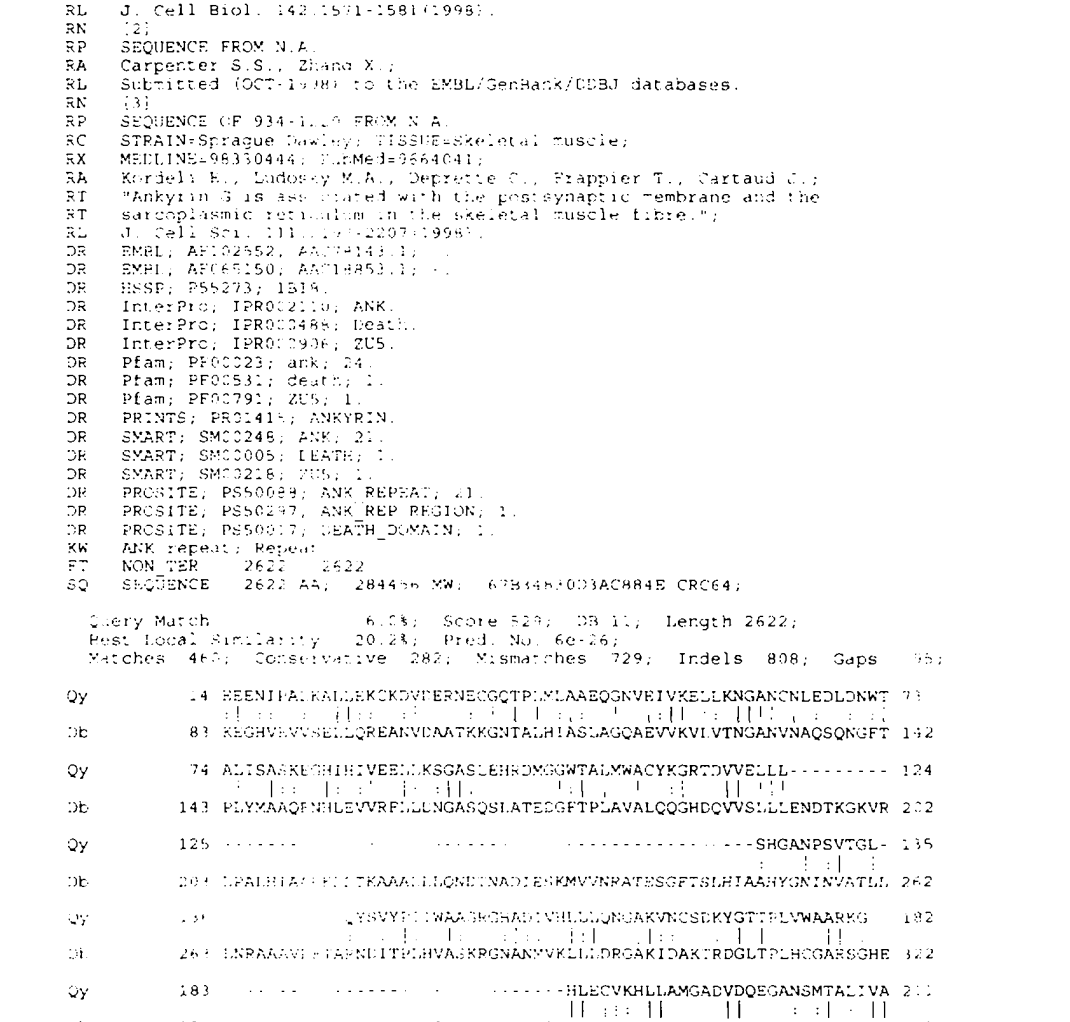
Query Match 6.2%; Score [illegible]; [illegible]
Best Local Similarity 30.6%; Pred. No. [illegible];

Matches 376; Conservative 270; Mismatches 604; Indels 574; Gaps 71;



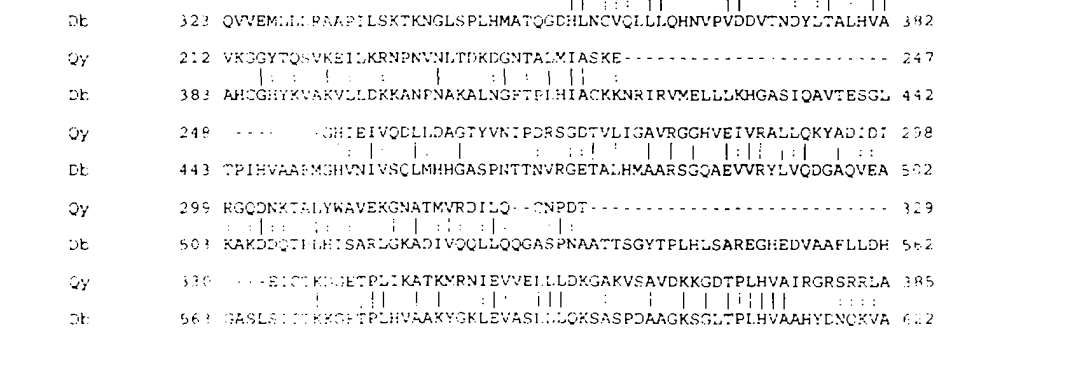


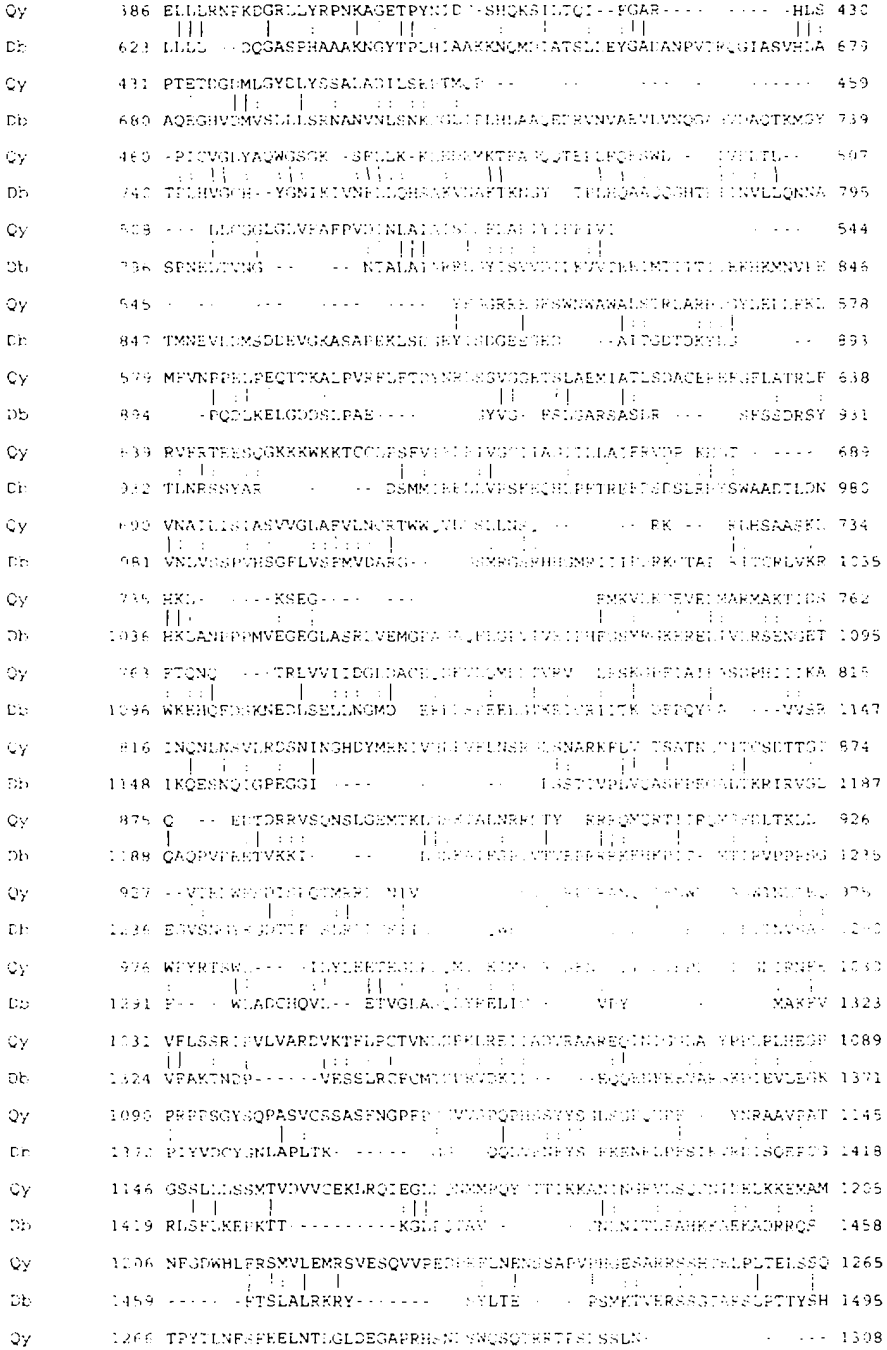
```
RESULT 15
Q70511
ID   Q70511      PRELIMINARY;      PRT,  2622 AA.
AC   Q70511;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 11, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   270 kDa ANKYRIN G isoform (ANKYRIN G)   Fragment.
GN   ANK3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98417633; PubMed=9744885;
RA   Zhang X., Bennett V.;
RT   "Restriction of 480/270-kD ankyrin G [illegible] requires multiple ankyrin G specific domains.";
```

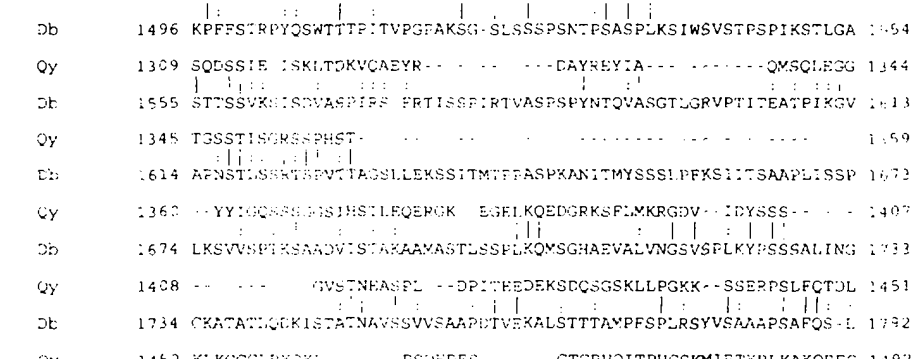


```
RL   J. Cell Biol. 142:1571-1581(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Carpenter S.S., Zhang X.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 934-1[illegible] FROM N.A.
RC   STRAIN=Sprague Dawley; TISSUE=Skeletal muscle;
RX   MEDLINE=98330444; PubMed=9664041;
RA   Kordeli E., Lambert M.A., Deprette C., Frappier T., Cartaud J.;
RT   "Ankyrin G is associated with the postsynaptic membrane and the
RT   sarcoplasmic reticulum in the skeletal muscle fibre.";
RL   J. Cell Sci. 111 ( Pt 15):2207-2218(1998).
DR   EMBL; AF102652; AAD19143.1; -.
DR   EMPL; AF065150; AAC18853.1; -.
DR   HSSP; P55273; 1SHA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000306; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
FT   NON_TER    2622    2622
SQ   SEQUENCE   2622 AA;  284456 MW;  67B648B1003AC884E CRC64;

  Query Match          6.0%;  Score 529;  DB 11;  Length 2622;
  Best Local Similarity   20.2%;  Pred. No. 6e-26;
  Matches  465;  Conservative  282;  Mismatches  729;  Indels  808;  Gaps   96;
```







Search completed: November 21, 2003, 12:47:06
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 12:[illegible] ; Search time 19.0 Seconds
without alignments
981.111 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 9883
Sequence: 1 MSVLISQSEVINYVEERNIFA[illegible]FIHAASNR[illegible]1716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 423108958 res.fues

Total number of hits satisfying chosen parameters: 3.[illegible]

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/ia/1A_COMB.pep:*
2: /cgn2_6/ptodata/1/ia/2A_COMB.pep:*
3: /cgn2_6/ptodata/1/ia/3A_COMB.pep:*
4: /cgn2_6/ptodata/1/ia/4A_COMB.pep:*
5: /cgn2_6/ptodata/1/ia/5A_COMB.pep:*
6: /cgn2_6/ptodata/1/ia/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	536	5.4	1839	2	US-09-[illegible]	Sequence 4, Appli
2	508.5	5.1	844	2	US-09-[illegible]	Sequence 3, Appli
3	466.5	4.7	1716	2	[illegible]	Sequence 1, Appli
4	466.5	4.7	1716	4	[illegible]	Sequence 1, Appli
5	466.5	4.7	1716	4	[illegible]	Sequence 1, Appli
6	466.5	4.7	1747	4	[illegible]	Sequence 4, Appli
7	466.5	4.7	1844	3	US-09-[illegible]	Sequence 2, Appli
8	414.5	4.2	1424	3	US-09-[illegible]	Sequence 1, Appli
9	385.5	3.9	1327	3	US-09-[illegible]	Sequence 2, Appli
10	385.5	3.9	1327	4	US-09-[illegible]	Sequence 2, Appli
11	394	3.9	1619	4	US-09-[illegible]	Sequence 4, Appli
12	392	3.9	1704	4	US-09-[illegible]	Sequence 2, Appli
13	380.5	3.8	949	3	US-09-[illegible]	Sequence 10, Appl
14	380.5	3.8	949	4	US-09-[illegible]	Sequence 10, Appl
15	376	3.8	348	2	US-09-[illegible]	Sequence 28, Appl
16	376	3.8	348	2	US-08-[illegible]	Sequence 29, Appl
17	376	3.8	348	3	US-09-[illegible]	Sequence 29, Appl
18	376	3.8	348	4	US-09-[illegible]	Sequence 28, Appl
19	373	3.8	1166	4	US-09-[illegible]	Sequence 5, Appli
20	359.5	3.6	352	3	US-09-[illegible]	Sequence 139, App
21	359.5	3.6	352	4	US-09-[illegible]	Sequence 139, App
22	357.5	3.6	741	2	US-08-[illegible]	Sequence 4, Appli
23	357.5	3.6	741	2	US-08-[illegible]	Sequence 4, Appli
24	357.5	3.6	741	2	US-08-[illegible]	Sequence 4, Appli
25	357.5	3.6	741	5	PCT-US95-[illegible]	Sequence 4, Appli
26	356	3.6	679	2	US-08-[illegible]	Sequence 4, Appli
27	356	3.6	679	2	US-08-[illegible]	Sequence 6, Appli
28	356	3.6	679	2	US-08-434-998-6	Sequence 6, Appli
29	356	3.6	679	2	US-08-487-797-6	Sequence 6, Appli
30	356	3.6	679	2	US-08-701-005A-4	Sequence 4, Appli
31	356	3.6	679	2	US-08-479-895-4	Sequence 4, Appli
32	356	3.6	679	5	PCT-US95-02058-6	Sequence 6, Appli
33	355	3.6	786	4	US-09-503-802-2	Sequence 2, Appli
34	355	3.6	787	3	US-09-188-930-334	Sequence 334, App
35	355	3.6	787	4	US-09-312-283C-334	Sequence 334, App
36	353	3.6	302	2	US-09-031-485-38	Sequence 38, Appl
37	353	3.6	302	2	US-08-847-429A-38	Sequence 38, Appl
38	353	3.6	302	3	US-09-065-474-38	Sequence 38, Appl
39	353	3.6	302	4	US-09-557-034-38	Sequence 38, Appl
40	353	3.6	303	2	US-09-031-485-23	Sequence 23, Appl
41	353	3.6	303	3	US-08-847-429A-23	Sequence 23, Appl
42	353	3.6	303	3	US-09-065-474-23	Sequence 23, Appl
43	353	3.6	303	4	US-09-557-034-23	Sequence 23, Appl
44	352	3.6	679	3	US-08-943-956A-4	Sequence 4, Appli
45	347.5	3.5	741	2	US-08-462-481-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF 0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10 14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: q29491
US-09-172-977-4

Query Match 5.1%; Score 536; DB 2; Length 1839;
Best Local Similarity 19.7%; Pred. No. 1.6e-35;
Matches 238; Conservative 2[illegible]; Mismatches 706; Indels [illegible]; Gaps 70;

Qy 14 EEENITALFALIEKGKPVDERNKDQQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWT 73
Db 73 KEGHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFT 132

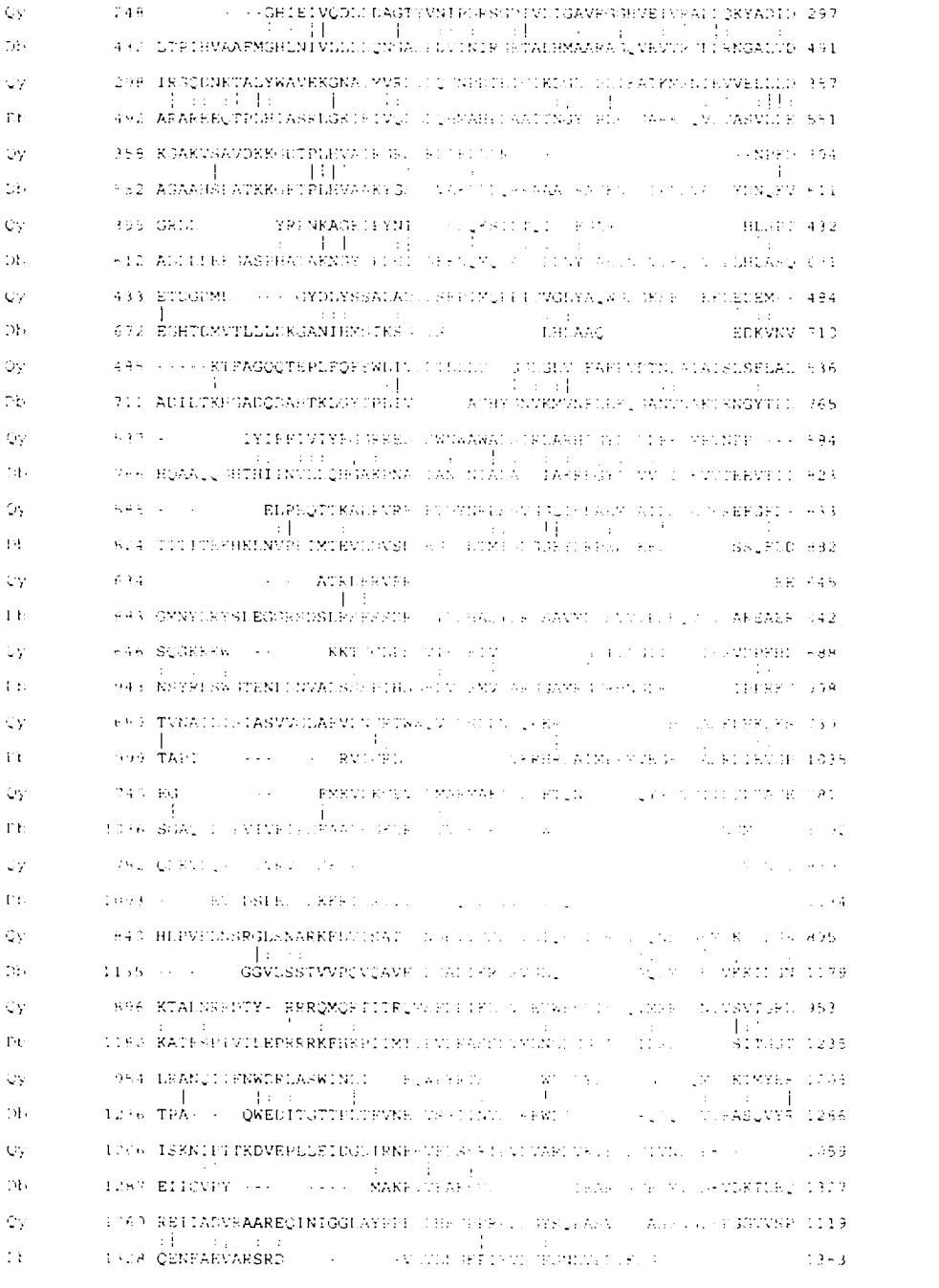
Qy 74 ALISASKEGHIEIVEELLKSGAS-- --- ----------LEHRDMG--- 103
Db 133 PLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVR 192

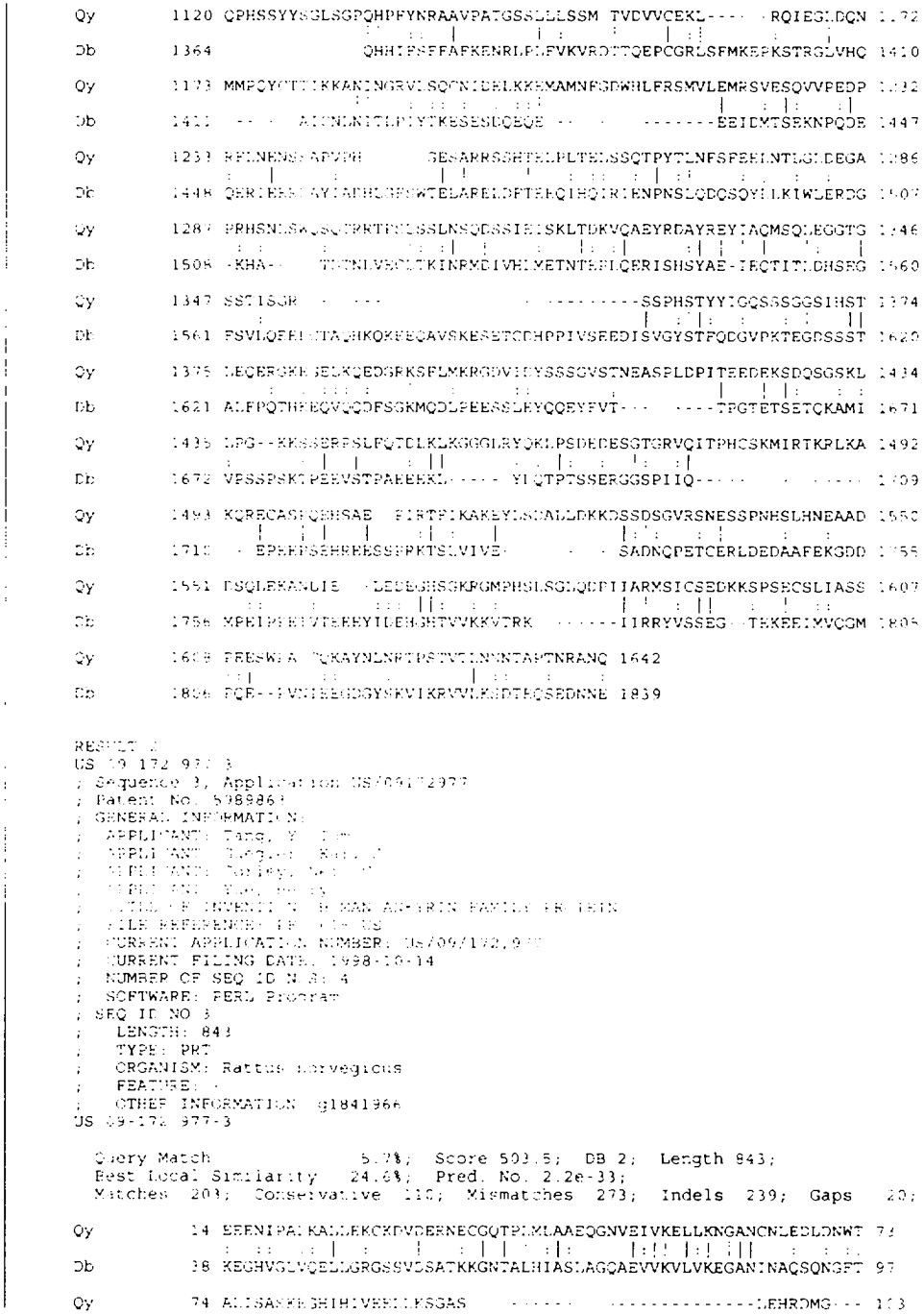
Qy 104 -- - - --- -- - ---- --------- --GWTALMWACYKGRTDVVELL 133
Db 193 LPALHIAARKNDTKSAALLLQNDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLL 252

Qy 134 LSHGANPSVTSLQYSVYPIIWAAGPGHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKG- 192
Db 253 LNRGAAVDFTA RNGITPLHVASKPGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 311

Qy 193 - --- -- ----- - -- -------HLECVKHLLAMGADVDQEGANSMTALIV 230
Db 312 DQVVELLLERGAPLLARTKNGLSPLHMAAQGDHVECVKHLLQHKAPVDEVTLDYLTALHV 371

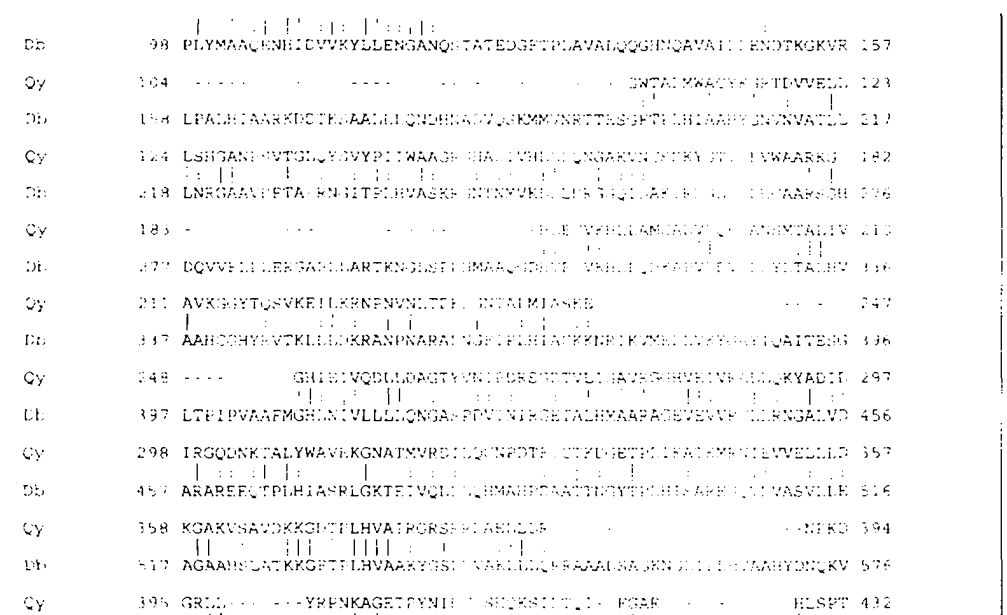
Qy 231 AVKGGYTQIVKEILKRNPNVNLTDKDGNTALMIASKE---- --------- ---- 267
Db 372 AAHCGHYEVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGASIQAITESG 431





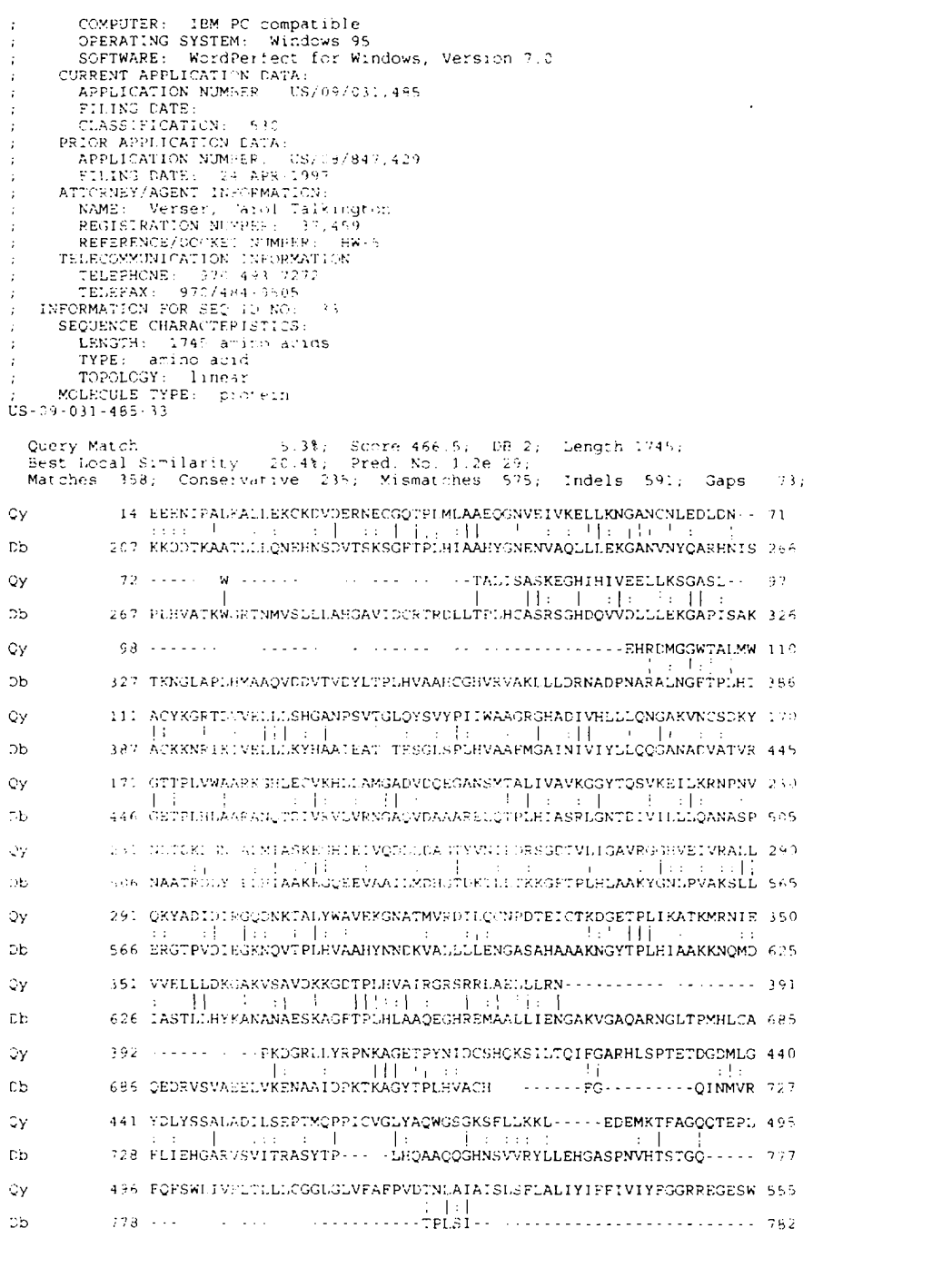
RESULT 3
US 09/172,977-3
; Sequence 3, Application US/09172977
; Patent No. 5988861
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. T.
; APPLICANT: Tang, Rui-J.
; APPLICANT: Turley, Sa.
; APPLICANT: Xu, Xiaoma
; TITLE OF INVENTION: A MAN ADRIN RECEPTOR
; FILE REFERENCE: 10.
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: gi1841966
US 09-172,977-3

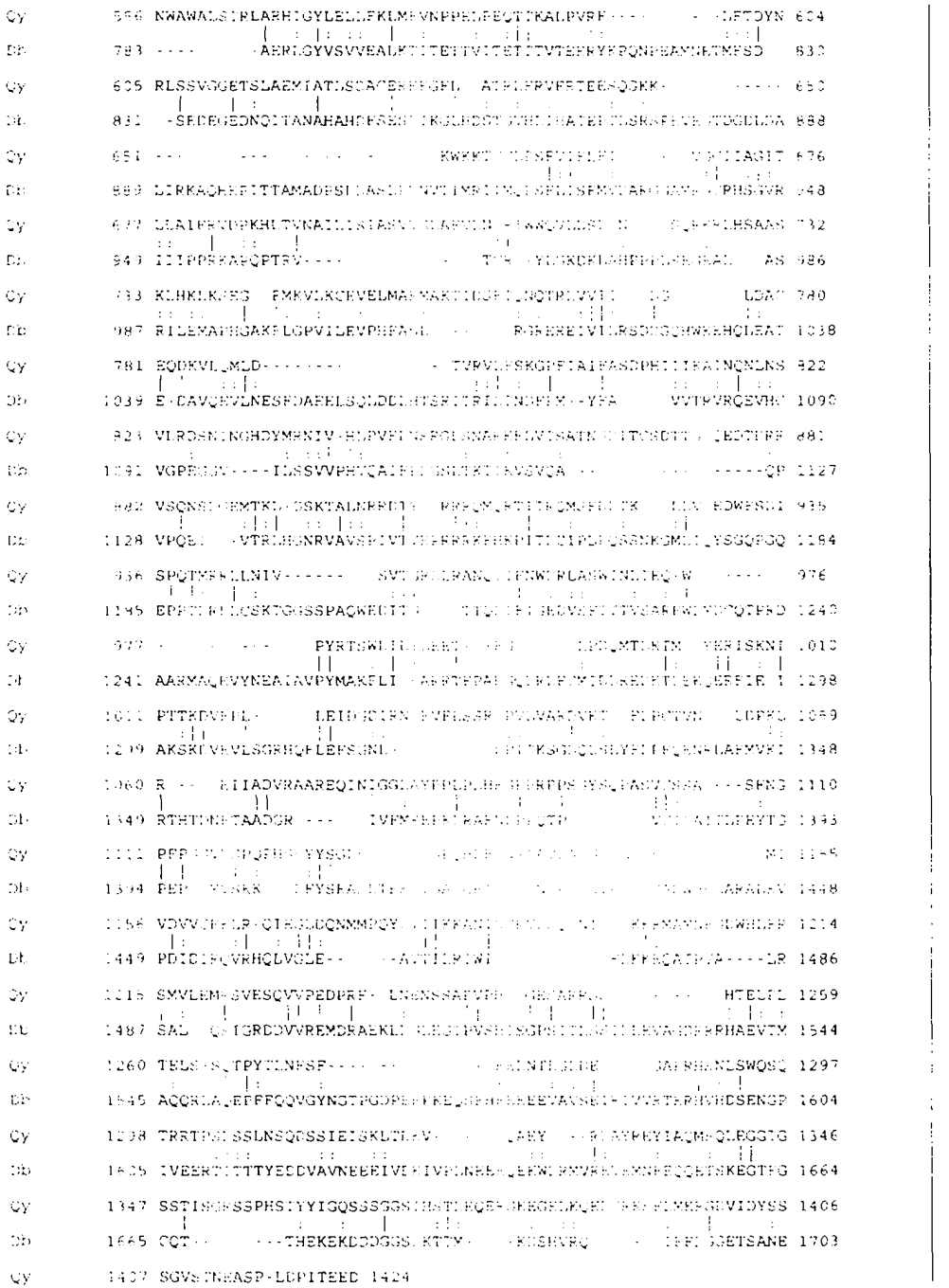
Query Match 5.7%; Score 503.5; DB 2; Length 843;
Best Local Similarity 24.6%; Pred. No. 2.2e-33;
Matches 203; Conservative 110; Mismatches 273; Indels 239; Gaps 20;

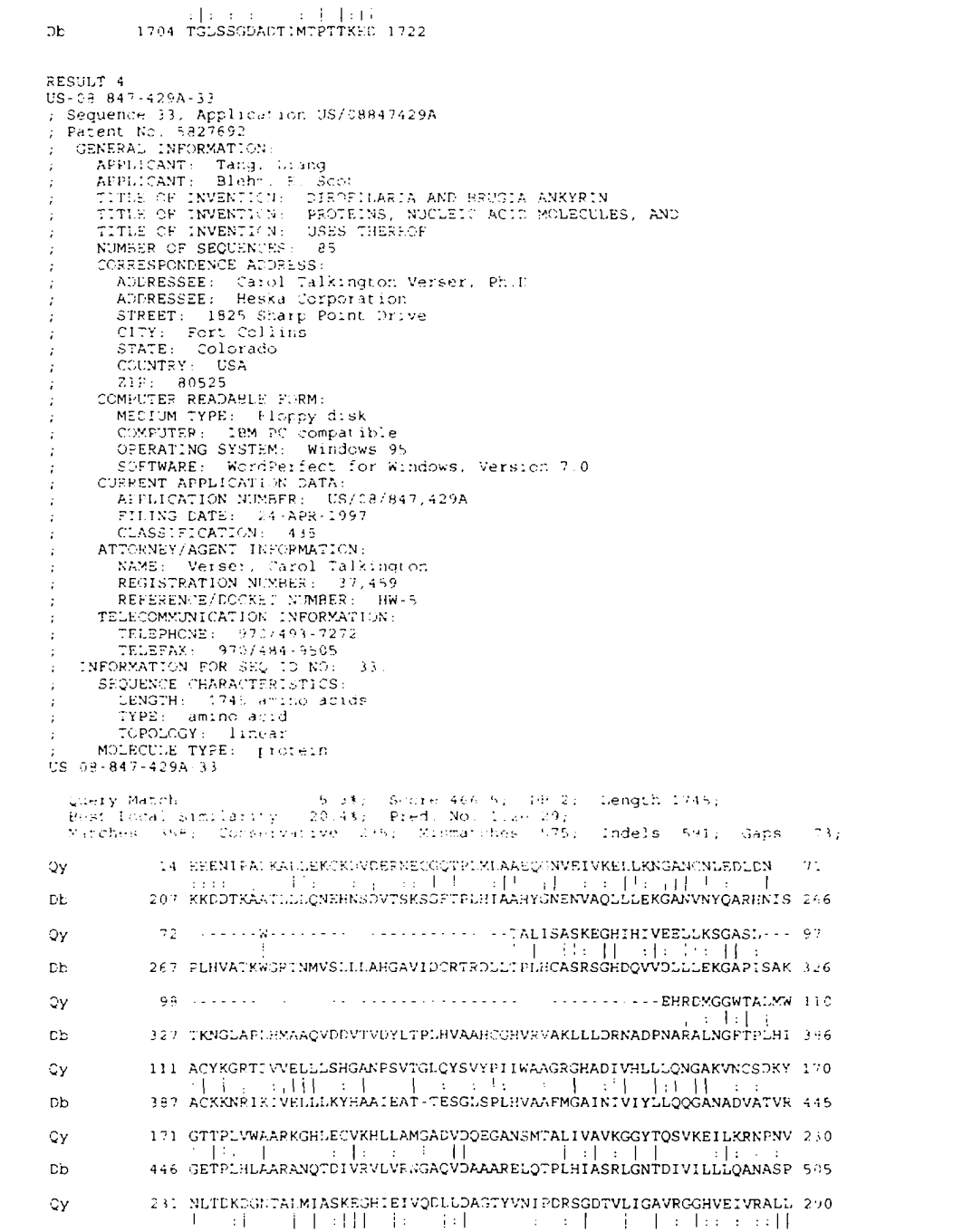


RESULT 3
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Bleay, E. Scot
; TITLE OF INVENTION: DIROFILARIA ANT BRUGIA ANTIGEN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Esq.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 38,469
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION
; TELEPHONE: 970 493 7272
; TELEFAX: 970/484-2505
; INFORMATION FOR SEQ ID NO: 33
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-33

Query Match 5.3%; Score 466.5; DB 2; Length 1745;
Best Local Similarity 20.4%; Pred. No. 1.2e-29;
Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;

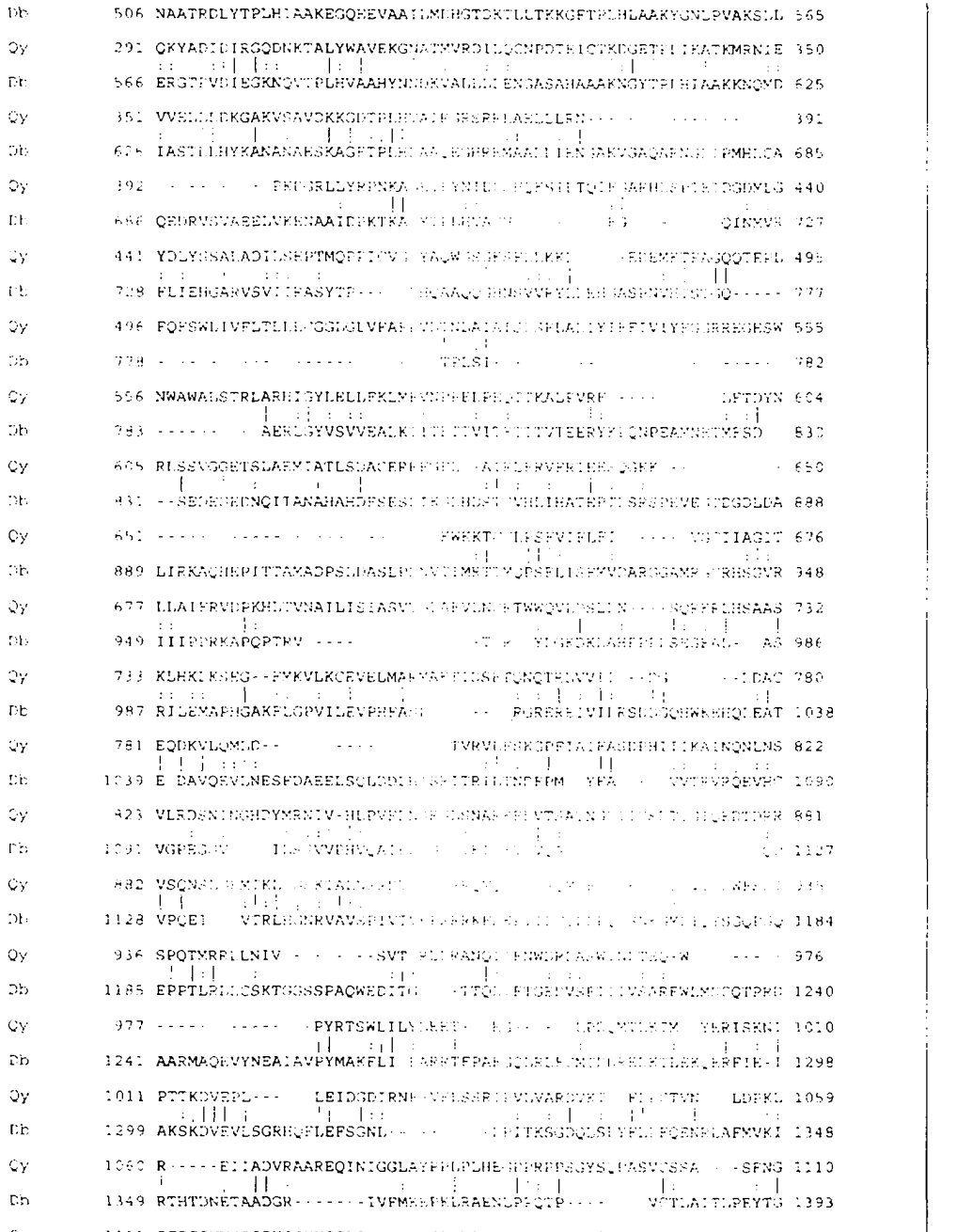


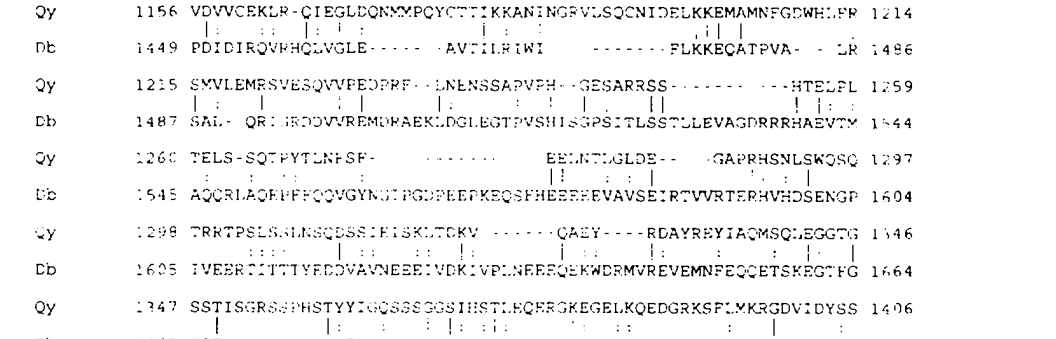




RESULT 4
US-08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, Kim Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,469
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-33

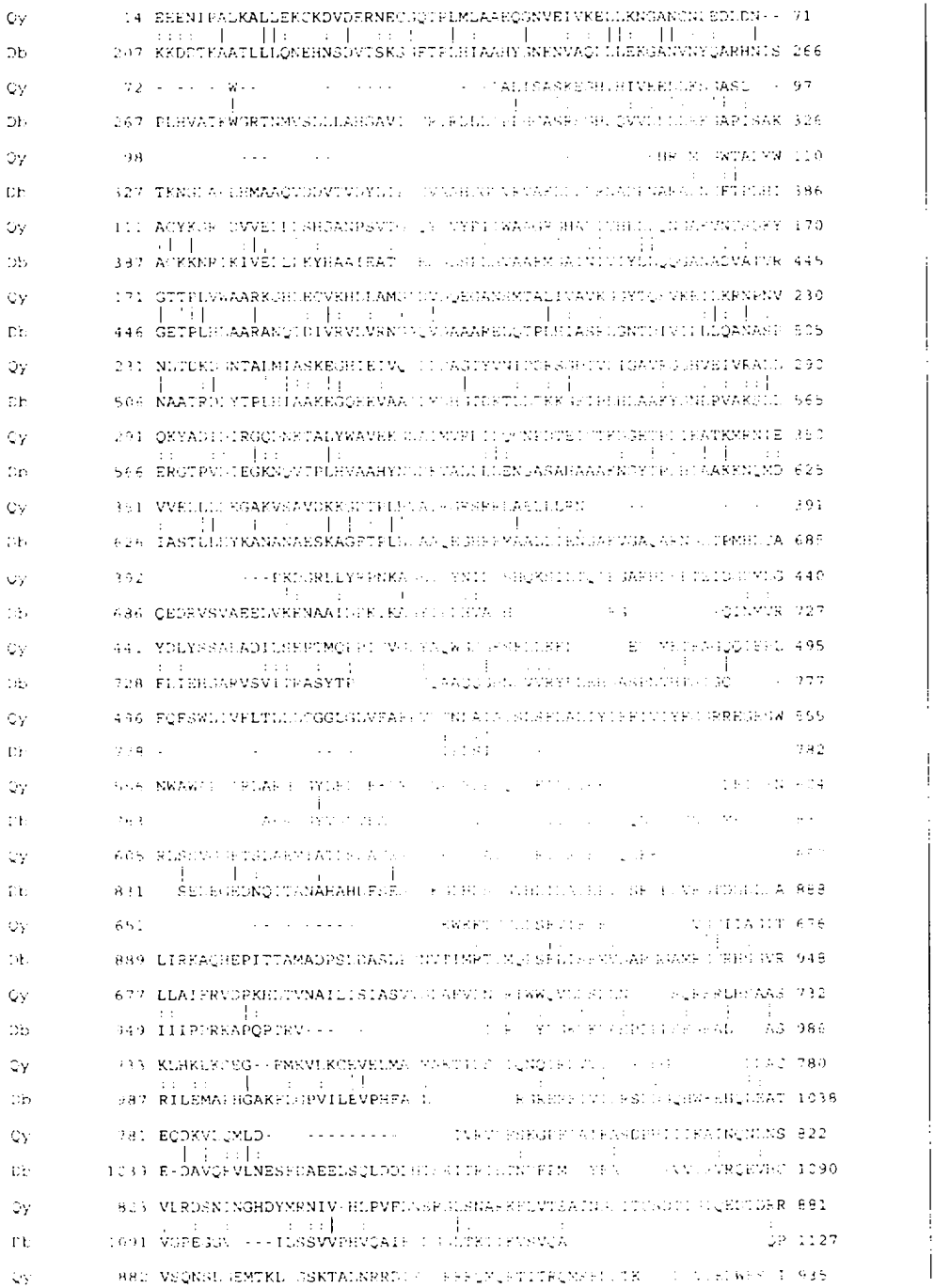
Query Match 5.2%; Score 466.5; DB 2; Length 1745;
Best Local Similarity 20.4%; Pred. No. 1.2e-29;
Matches 358; Conservative 276; Mismatches 575; Indels 541; Gaps 73;

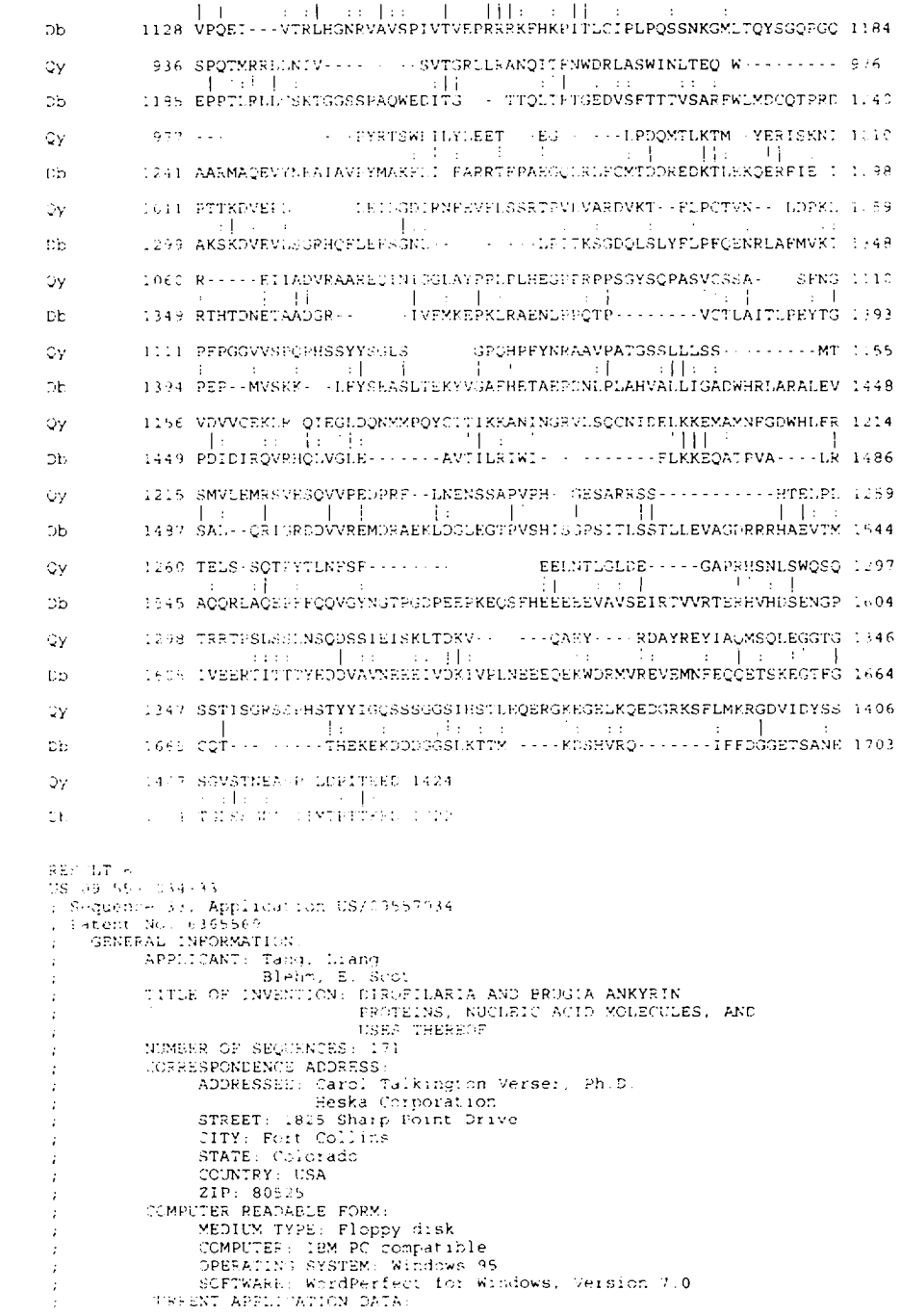




RESULT 5
US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, R. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-33

Query Match 5.3%; Score 466.5; DB 3; Length 1745;
Best Local Similarity 20.4%; Pred. No. 1.2e-29;
Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;





```
RESULT 6
US/09/55[illegible]334-33
; Sequence 33, Application US/09557934
; Patent No. 6365569
;   GENERAL INFORMATION:
;       APPLICANT: Tabai, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
```

```
;            APPLICATION NUMBER: US/09/547,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/075,424
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Ta[illegible]
;            REGISTRATION NUMBER: 37,[illegible]
;            REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION
;            TELEPHONE: 970/493-[illegible]
;            TELEFAX: 970/484-95[illegible]
;   INFORMATION FOR SEQ ID NO: 33:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1745 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US 09/547,034 33

  Query Match              5.3%;  Score 466.5;  DB 4;  Length 1745;
  Best Local Similarity   20.4%;  Pred. No. 1.1e-20;
  Matches  358;  Conservative  235;  Mismatches  575;  Indels  197;  Gaps  73;

Qy          14 EEENIPALKALLEKIKDVDEENKYTLMIAAEQINVEIYNFLIKNIDADIMIEDLCN   71
               :::: |  :|:  : | :: :  :|| . : . :|:  :| :  :
Db         207 KKDLIKAATQLLQNEHNSDVTSKSHTTIHIAAHYCNENVAQILLEKGAQVNSQARNNIS 266

Qy          72 --      W ---- :  ----      -- TALISAARSGHIEIVEELLKQGASL--  97
                        :      |    |:  :  | : |  |  :
Db         267 PLHVATFWGRTNMVSLLLAIGAVI TRPLLLDITDAFRKGH QVVDILIEFGAPISAK 326

Qy          98  - -  -  -  ---- --  -                    MRPIHIAVTALMW 130
                                                        |  : : :
Db         327 TKNDIAPLHMAAQVPDVTVILLTPLHVAAHFGIVAVAELLIRVGAAINAPALVFPTLHE 386

Qy         131 ACYKGRTEVVELLISHGANPSVTGLNYVTPLHMAAERGHAEIVSLLLKLNAEVNCSDKY 190
               |' :  :  :|||:  : |  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         387 ACKKNRIKIVELLLKYHAAIFAT-TKSG-SPLHLAAAMGAIETIEVLQKHGADVATVD 445

Qy         191 GTTPIVWAARKGHLEGVKHLIAMGADPVTLRGANIMALNVAEFYTQASVERLKKNSDV 250
               | :  |||    :  : :  ||  :  :  :  :  :  :  :  :  :  :  :  :
Db         446 GETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIAGRLGTDIVELLQANAGP 505

Qy         251 NLDTRDIMGATALMIAKEGHTEIVQELLKRGVDVNIKNKVGWTPLMVAKTQNHLEIVK 310
               |  :  :  :  :  :|  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         506 NAATRSGFTPLHLSAQQGHTDIVELLLSKQASVNAVTKKSFTPLHLAAQEGHVEVVE 565

Qy         311 QKYALITRPSQDAIASNGHKEIKEILQNREVKHSRSRSPSSSSSSSSSSSSSSSSSSSS 370
               ::  :|  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         566 ERGTPVDIEGKNQVTPLHVAAKYNSPKVAQLLLSKGASPNLQNVSGLTPLHLAARLGY 625

Qy         371 VVELLLSKGAKVSAVDKKGITPLHLAIQKGHE AAIIEI                    -  430
               :  :  |  :  :  :|||:  :  :  :  :  :  :  :  :  :  :  :  :
Db         626 IASTLLHYKANANAESKAGFTPLHIAAQEGHTEVAALLEVNGFSIQARTKLGYTPMHLA 685

Qy         431 -  --  -  ---PKDGRLLYPPNKA-IIIYNII  HQKSIIQQI  APHLSPTHIGGIMI 489
                        |:  :  :  '|'  :  :  :  :  :  :  :  :  :  :  :  :
Db         686 QEDRVSVAEELVKENAAIDPKTKA-GLTPLHVASF                QIDVVH 727

Qy         490 YDLYSSALADILSEPTMQPPITVQLAQAVPAGWEEIARTASGQWFFVDMKSIRYEDSIR 549
               :  :  |  :::  :  |  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         728 PLIEHGARVSVITRASYTP    -E QAAQQ RDDTVRVLLEHGADVNAQDKFGKTAFQ --  777

Qy         550 FQFSWLIVPLTLLLMGGLGLVFAFVQDLAIA CSEIAIEA IPIGIYPIYGWREIESW 609
                                    :|                                :
Db         778 ---      ---      --  --    |||SI                    --   782

Qy         610 NWAWALSTRLARHIGYLELLFKIMEKNPTELPPIDLKAAIPVAF            IFIDYN 664
               :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         783 -  --  --AERDIGYVSVVEALKTHPTIVIDITPVIEERPEFQNPVAMALVMRGD  840

Qy         665 RLSSVGGETSLAEMIATLSDADREFGFL--ATRLFRVFRTEESQGKK------- -- - 710
               |  || :  :  | '  '  :| :  :  |: :  :
Db         841 --SEDEGEDNQITANAHAHDFSESLTKGLHDSTGVHLIHATEPTLSRSPEVEGTDGDLDA 898

Qy         711 - -  -  -  --------- ----KWKKTCCLPSFVIFLFI-------VGCIIAGIT 676
                                    :|  |||:|  :  :|  :|:
Db         899 LIRKAQHEPITTAMADPSLDASLPDNVTIMRTTMQPSFLISFMVDARGGAMRGCRHSGVR 958

Qy         677 ILAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLLN----SQRKRLHSAAS 732
               :|  :  |  :  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         959 IIIPFRKAPQPTRV--  ------  -----TCP- YLGKDKLAHPPPLSEGEAL- -AS 996

Qy         733 KLHKKLKSEG -PMKVLKCEVELMARMAKTIDSPTQNQTRLVVII---DG  ----LDAC 780
               :: ::  |  :|:  :  ||  |  :  :  :  :|:  :  |  :|:
Db         997 PILEMAPPGAKFLGPVILEVPHFASL  -  ---RGRREEIVILRSDDGQHWKEHQLEAT 1038

Qy         781 EQDKVLQMLD--------  -  ----TVRVLFSKGPFIAIFASDPHIIIKAINQNLNS 822
               :|  ::|:                  |:  :  |  ||  :  :  :|:  :
Db        1039 E-DAVQEVLNESPDAEELSQLDDLHTSRITRLTNDFPM--YFA-----VVTRVRQEVHC 1090

Qy         823 VLRDSNINGHDYMRNIV-HLPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRR 881
               :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1091 VGPEGGV--- ILSSVVPHVQAIFPDGSITKTIKVSVQA ---------  -  --QP 1127

Qy         882 VSQNSLGEMFKL-GSKTALNRRDTY- RRFQMQRTITRQMSFDLTK---LLVTEDWFSDI 935
                  |  :  :| |:  :|:  |  |:|:  :  :|  :  :  :|
Db        1128 VPQSI- -VTRLHGNRVAVSPIVTVEPRRRKFHKPITLCIPLPQSSNKGMLTQYSGQPGQ 1184

Qy         936 SPQTMRRLLNIV---------SVTGRLLRANQITFNWDRLASWINLTEQ W--------- 976
                  | |:  | :          :||      |:|:  :  :  :  :
Db        1185 EPPTLPLLCSKIGGSSPAQWEDITG---TTQLTFTGEDVSPTTTVSARFWLMDCQTPRD 1240

Qy         977 ------ - ------PYRTSWLILYLEET  -EG------LPDQMTLKTM--YERISKNI 1019
                     :  :  :  :  :  :  :  :  ||      :  :  ||:  ||  :
Db        1241 AARMAQEVYNEAIAVPYMAKFLI-PARRTFPAEGQIRLRFCMTDDRENDKTLEKQERFIE 1299

Qy         1011 PTTKDVEPL------LEIDGDIRNFEVPLSSRTPVLVARDVKT--FLPCTVN---LDPKL 1069
                  :|:  |      ,  |::      |:  :  |  :  |||          :
Db        1299 AKSKDVEVLSGRHQFLEFSGNL  -  -     LPITKSGDQLSLYFLPFQENRLAFMVKI 1358

Qy         1070 P-    EIIADVRAAREQINIGGLAYPPLPLHEGPPRPPSGYSQPASVCSSA----SFNG 1119
                  :  |  |       |  :  |  :  :  |  :|      ||:      :  :
Db        1359 RTHTDNETAADGR--      IVFMKEPKLRAENLPFQTP--------VCTLAITLPEYTG 1398

Qy         1120 PFPGSVVQPQTHSSYYSQLS--  ---GPQHFFYNRAAVPATGSSLLLSS---------MT 1165
                  |  :  :  :  :  :  |          :|    :  :  :  :
Db        1399 PEP  PVPEK   LFYDRAALTEKYVGAFHETAEIMNLPLAHVALLIGADWHRLARALEV 1458

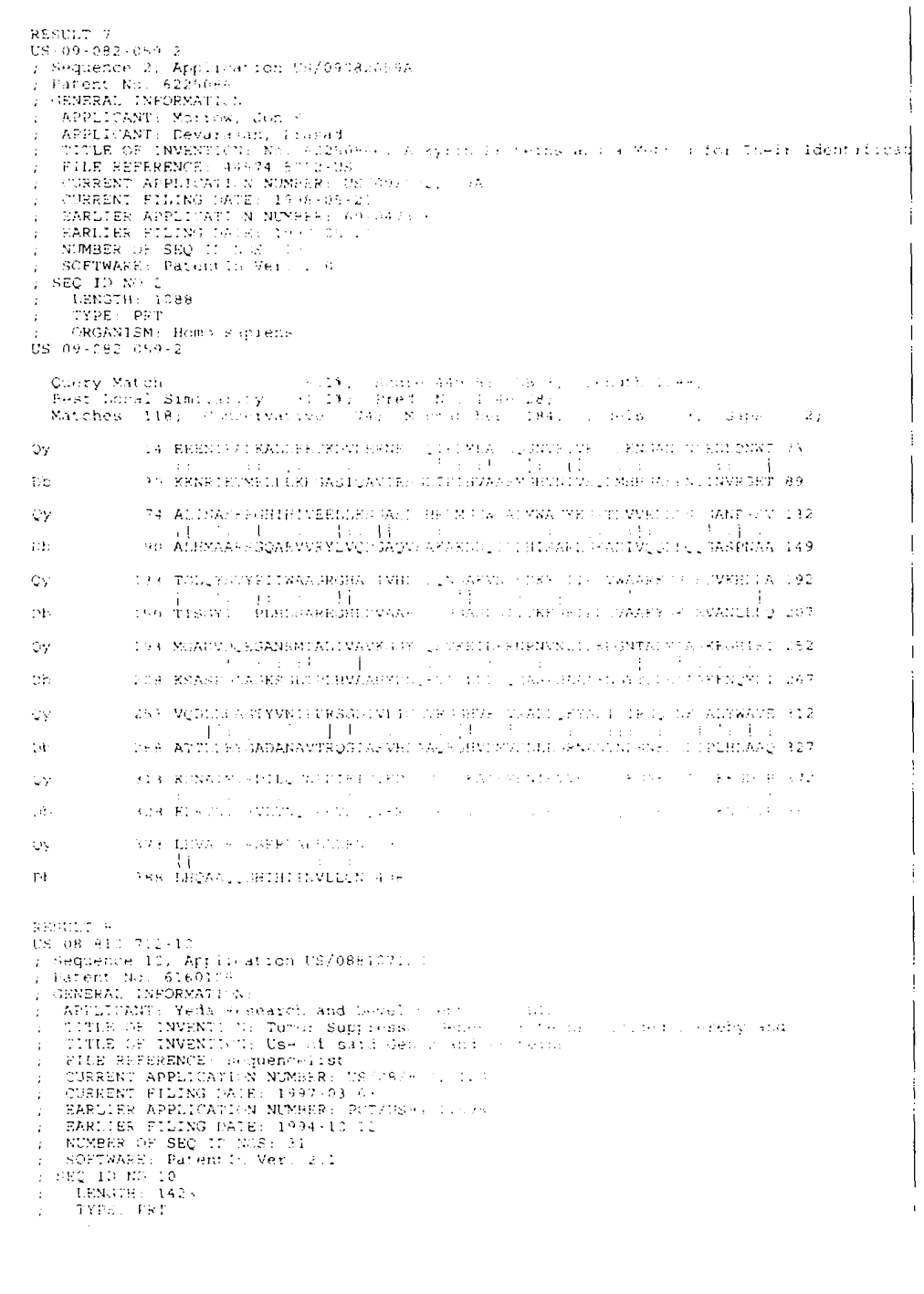
Qy         1166 SMVLEMPSVESQVVPELPRP- LNENSSAPVPH--GESARRSS----------HTELPL 1219
                  :  :|  |  :  :  :|  :  |  :  :|  :  :  :|  :  :  :  :  :
Db        1459 SAL--QRIGRDDVVREMDRAEKLDGLEGTPVSHISGPSITLSSTLLEVAGDRRRHAEVTM 1516

Qy         1220 IELS SQTPYTLNFSF--  ------------EELNTLGLDE-----GAPRHSNLSWQSQ 1267
                  :  :| |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1517 AQQRIAQRPFFQQVGYNGTPGDPEEPKEQSPHEEEEEVAVSEIRTVVRTERHVHDSENGP 1576

Qy         1268 TRRTPSLSSLNSQDSSIEISKLTDKV-------QAEY----RDAYREYIAQMSQLEGGTG 1316
                  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1577 IVEERTITTTYEDDVAVNEEEIVDKIVPLNEEEQEKWDRMVREVEMNFEQQETSKEGTPG 1636

Qy         1317 SSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERGKEGELKQEDGRKSFLMKRGDVIDYSS 1376
                  |          |:  :  :  ||:  :  :  |:  ::  :  :  :  :
Db        1637 CQT--      --  THEKEKDDGGSLKTTM- ---KDSHVRQ-------IFFDGGETSANE 1703

Qy         1377 SGVSTNEASP-LDPITEED 1424
                  :|:|:  |  :  | :: |
Db        1704 TGLSSGDADTIMTPTTKED 1722
```


RESULT 7
; SEQ ID NO: 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
RESULT 9
; ORGANISM: Homo sapiens

```
Db       1056 RYTVEDIQRLVPDSDVEELLQILDAMDICARDLSSDTM-  --  VDVPA-LIKTENLHRS 1108

Qy        963 W 963
              |
Db       1109 W 1109


RESULT 9
US-09-196-387-2
; Sequence 2, Application US/09196387
; Patent No. 6277617
; GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,229
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  no
US-09-196-387-2

  Query Match             4.8%;  Score 449.5;  DB 2;  Length 1327;
  Best Local Similarity   23.3%;  Pred. No. 9.4e-44;
  Matches  167;  Conservative   80;  Mismatches  211;  Indels  21;  Gaps   17;

Qy         14 EEENIPALKALLEKCKDVDERNECGLTPLMLAAL--NVEIVEFLIKHGANVNLKDLNWT 71
              |:: :  |  :  |  | |    ||   ||   | | : ||   |: : :|
Db        348 EEKLMALLTPLNVNCHASDGRK--NTLIHLASINNRVRIVLSLQHGADVHAADKQGLV 404

Qy         74 ALISASKEGHIHIVEELLKSGASLHHTMQNWTALYWAIYEHKTAVVELLISHGANPSV- 132
              |:   ||  : | | |  ||  :  :   :  :  :  |  :  :  ,|| :|::
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASFNKAEVCSLLLSLGADPTLV 464

Qy        133 --- --- ---------- -------  --       -  -  ---  -  ----T 133
                                                                    |
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHQLLQAAREADIAKVKKILALEIINFKVPQSHET 524

Qy        134 GLQYSV-------  -------  ----  --   -YPIIWAAGFGHADIVHLLLQNGAKV 164
                 |  :                        :  :| | ' |:  :| :|:||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNSKNKDFMTPLHVAARFAHNDVMVLLKHGAKV 584
```

```
Qy        165 NCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTA------ ------ --- 207
              |  :  |   |  ||  :|:  :  |: |:|          ||
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        208 -----  ---LIVAVKGGYTQSVKEILKRNPNVNLTDKDG--NTALMIASKEGHIEIVQDL 256
                     |: | | '  :: :  :  : |||    :|  :| |  |:    : :|: |
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCS-SQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYL 703

Qy        257 LDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKG-- 314
              |    |: |:  |  |  ,  '' |:   |:: |  :::    | ': |  ||
Db        704 LHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKY 763

Qy        315 - - --- --  -----   -- --  --NATMVRDILQ----------------  -C 325
                                          | : :|:                    |
Db        764 EICKLLLEHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLC 823

Qy        326 NPDTEICTKD  -GETPLIKATKMRNIEVVELLLDKGAKVSAVDKKGDTPLH-------- 374
                |:   |        |||  :      :||  | |  | || |:| |  |'  ||
Db        824 TPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHV 883

Qy        375 ----- -- ---- --- --- --- VAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETP 408
                                        | :|: :   |||  :  |  :    |: :::
Db        884 DIAALLLKYNTCVNATDKWAFTPLHEAAQKGRT-QICALLLAHGADPTM-  KNQEGQTP 939

Qy        409 YN------ -  -IDCSHQKSILT------QIFGARHLSPTETDGDMLGYDLYSSALADI 451
                 :          ||  :::  |      :  |  :||  |       | :::   |
Db        940 LDLATADDIRALLIDAMPPEALPTCFKPQATVVSASLISPASTPS-----GLSAASSIDN 994

Qy        452 LSEPTMQFFICVGLYAQWGSG------  ------ -------KSFLLKKLEDEMKT 486
                '| |  :  :  :  :  |                        |:  :  :|
Db        995 LTGPLAE- LAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLFHLRDIFET 1049


RESULT 10
US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. 6506587
; GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-841-835-2

  Query Match             4.5%;  Score [illegible];  DB 4;  Length [illegible];
  Best Local Similarity   25.3%;  Pred. No. [illegible];
  Matches  167;  Conservative   [illegible];  Mismatches [illegible];  Indels   [illegible];  Gaps   17;

Qy          14 EERNIPALKALIEKFKDVERNEQ[illegible]MLAARL RVRIIVPPLSRDIAL TFEDIINAT 73
Db         348 EEKLMALLTPLGVVNCHASDGRK- [illegible] 404

Qy          74 ALISASKEGHIHIVEELLKNGASLFIR NRDWTALHWA TYNGHTDVVEL NEGANPSV 132
Db         405 PLHNACSYGHYEVTELLIKHGACVNAMD WQFTPLHEAA KNGYVDIA SVGADPTLV 464

Qy         133 [illegible] 193

Db         465 NCHGKSAVDMAPTPELRELLTYPPF [illegible] 524

Qy         134 GLQYSV [illegible] 164

Db         525 ALHCAVASLHPKRKQVTELLLKKGANVNLKNKL [illegible] 584

Qy         165 NCSDKYTPLHVAAARKGHLECVKL [illegible] 207

Db         585 NAIDTTGQTALHRAALAGHLQTCR [illegible] 644

Qy         208 [illegible] 256

Db         645 IRTSIVDYRLLEAKKAGDLETVKC [illegible] 703

Qy         257 LDAGTYVNIPDRGGDTVLIAVRGQ [illegible] 314

Db         704 LHHGADVHAKDKGGLVPLHNACSY [illegible] 763

Qy         315 [illegible] 375

Db         764 EICKLLLKHGADPTKKNRDGNTPI [illegible] 823

Qy         376 [illegible] 434

Db         824 TPENIMYKDTQGRNTLHAAGYSN [illegible] 883

Qy         435 [illegible] 477

Db         884 DIAAVLKYNTMTVLELFWAFTL [illegible] 943

Qy         478 YN [illegible] 491

Db         944 LDLATADNIRALLIDAMPTEALPT [illegible] 994

Qy         492 LSEPTMQPPICVGLYAQWGSG [illegible] 546

Db         995 LTGPIAE LAVGGASNAGDGAA [illegible] 1049


RESULT 11
US 09-392-812A-4
; Sequence 4, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
;  APPLICANT:  Zuker, Charles S.
;  APPLICANT:  Walker, Richard G.
;  APPLICANT:  Willingham, Aarron
;  APPLICANT:  The Regents of the Univer[illegible]ty of California
;  TITLE OF INVENTION:  A Eukaryotic M[illegible]anosensory Transduction Channel
;  FILE REFERENCE:  [illegible]307E-097600US
;  CURRENT APPLICATION NUMBER: US/09/[illegible]A
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 4
;  LENGTH: 1619
;  TYPE: PRT
;  ORGANISM: Drosophila melanogaster
;  FEATURE:
;  OTHER INFORMATION: amino acid sequence derived from nompC cDNA
;  OTHER INFORMATION: sequence
US-09-392-812A-4

  Query Match             4.5%;  Score 194;  DB 4;  Length 1619;
  Best Local Similarity   25.2%;  Pred. No. 1.6e-23;
  Matches  149;  Conservative   98;  Mismatches  199;  Indels  146;  Gaps   21;

Qy          29 KDVDERNEQGQTPLMLAAEQGNVEIVKELLDKNGANCNLEDLDNWTALISASKEGHIHIVE 88
Db         651 KAMNRQSSVGWTPLLIACHRGHMELVNNLLANHARVDVFDTEGRSALHLAAERGYLHVCD 710

Qy          89 ELLKSGASLERRDMQGWTALMWACYKGRIDVVELLL-SHGANPSVTGLQYSVYPIIWAAG 147
Db         711 ALLTNKAPINSKSRVGRTALHLAAMNGFTHLVKFLIKDHNAVIDILTLRKQT-PLHLAAA 769

Qy         148 RGHADIVHLLLQNGAKVNCSDKYGITPLVWAARKGHLECVKHLLAMGADVDQEGANSMTA 207
Db         770 SGQMEVVQLLLELGANIDATDDLGQKPI- ---  ----H-------- ------ -- 798

Qy         208 LIVAVKGGYTQSVKEILKRNPN-VNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIP 266
Db         799  -VAAGNNYSEVAKLFLQQHPSLVNATSEDGNTCAHLAAMQGSVKVIELLMKF- ---  849

Qy         267 DRSG- --  -DTVLIGAVRGGHVEIVRALLQKYAD---------- -- ---  295
Db         850 DRSGVISARNKLTDATIQLAAEGGHADVVKALVPAGASCTEENKAGFTAVHLAAQNGHG 909

Qy         296  -IDIRGQDNK  --  TALYWAVEKGNATMVRDILQCNPDT---  ----- E 330
Db         910 QVLDVLKSTNSLRINSKKLSLTPLHVAAYYGQADTVRELLTSVPATVKSETPTGQSLFGD 969

Qy         331 ICTKDGETPLIKATKMRNIEVVELLLQK GAKVSAVD-KKGDTPLHVAIRGRSPRLAELL 388
Db         970 LGTESGMTPLHLAAFSGNENVVRLLLNSAGVQVDAATIENGYNPLHLACFGGHMSVVGLL 1029

Qy         389 LRNPKDGRILYRPNKAGETPYNIDCSHQKSILTQIP---GARHLSPTETDGDMLGYDLYS 445
Db        1030 L  SRSAELLQSQDRNGRTGLHIAAMHGHIQMVELLGQGA-EINATDRNG---WTPLHC 1083

Qy         446 -ALA  [illegible]  [illegible] 494

Db        1084 AAKAGHLEVVELLCEA-GASPKSETNYGCAAIWFAA-EGHNEVLRYLMNKEHDTYGLMEDK 1143

Qy         495 LFQFSWLIV [illegible] 530

Db        1144 RFVYNLMVVSKNHNNKPIQEFV------- -LVSPAPVDTAAKLSNIYIVLS 1186


RESULT 12
US 09-392-812A-2
; Sequence 2, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
;  APPLICANT:  Zuker, Charles S.
;  APPLICANT:  Walker, Richard G.
;  APPLICANT:  Willingham, Aarron
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION:  A Eukaryotic Mechanosensory Transduction Channel
;  FILE REFERENCE:  [illegible]307E-097600US
;  CURRENT APPLICATION NUMBER: US/09/392,812A
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
```

```
;   LENGTH: 1704
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence derived from [illegible]
;   OTHER INFORMATION: sequence
US-09-392-812A-2

  Query Match             4.4%;  Score [illegible]
  Best Local Similarity   [illegible]%;  Pred. No. [illegible]
  Matches  231;  Conservative  157;  Mismatches [illegible];  Indels [illegible];  Gaps  48;

Qy         29 KDVDERNEGQTPLMLAAEQGNVE[illegible] 88
Db        316 KAMNKQSSVGWTPL[illegible] 375

Qy         89 ELLKSGASLEHRDMGGWTALMWA[illegible] 147
Db        376 ALDTNKAFINSKSRVGRTALHLAAMN[illegible]PLHLAAA 434

Qy        148 RGHADIVHLLLQNGAKVNCSDKYGT[illegible] 207
Db        435 SGQMEVCQLLLELGANIDATDDLGQ[illegible] 463

Qy        208 LIVAVKGGYTQSVKEILKRNPN-[illegible]GTYTNIP 266
Db        464 --VAAQNNYSEVAKLFLQQHPSLVNA[illegible] 514

Qy        267 DRSG---------DTVLIGAVRGG[illegible]KYAT 295
Db        515 DRSGVISARNKLTPAIPLQLAAEG[illegible] 574

Qy        296 --IDIRGQNK-[illegible] 330
Db        575 QVLDVLFSTNSLRINSKKL[illegible] 634

Qy        331 ICTKDGETPL----[illegible] 364
Db        635 LGTESGMTPLHLAAFSGNENVVRLL[illegible] 694

Qy        365 VDKKGDTPLHVAIRGRSRRLAELL[illegible]LLYEPNK 403
Db        695 TDRNGWTPLHCAAKAGHLEVVKLL[illegible] 754

Qy        404 AGET[illegible]
Db        755 [illegible]

Qy        452 LSE[illegible]YAQW-[illegible] 488
Db        814 AKEIVAA[illegible]KQCLAMATELLALAA[illegible] 873

Qy        489 GQQTEPLEQ----FSW--LIVFLT[illegible] 541
Db        874 QRYLQELWHGSLTWASWKI[illegible] 928

Qy        542 IVIY[illegible]EGESWNWAWALSTRLA[illegible] 601
Db        929 [illegible] 951

Qy        602 DYNRLS-----SVGGETSLAEMIA[illegible] 655
Db        952 --RLSLVPYWYEVG-------LLIW[illegible] 984

Qy        656 CCLPSFVIFLFIVGIIAGITLLA[illegible] 709
Db        985 --LGSIKVLVLLLGMAGVGVHVSAF[illegible] 1040

Qy        710 -----RTWWQVLDSLLNSQRK[illegible]RLHSAA[illegible]SRLHK 736
Db       1041 LSFHHLFGPWAIIIGDLLKDLARF[illegible]RSFEKF 1100

Qy        737 LKSEGF--MKVLKCEVELMARMA[illegible] 793
Db       1101 NRNRGYFSDMEQMTCPHPDLRRW---------RIMSIVASANSDE--  ---  -STR 1139

Qy        794 VLFSKGPFIAIFASDPH--  ---  ---  -IIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Db       1140 TTFPGGT  ---  STSPHSLLEIPSPQMHVDVFIQSIQTKIKQSISNIDITN --   --AR 1187

Qy        841 LPVFLNSR[illegible]SNARKFLVTHATNGD-    ITCSDITGTQERTDRRVSQNSL[illegible]EMTK-  --  892
Db       1188 LPGAFSLRRLPTT KF[illegible]IETDFIESITKNENA - TDTDYRCSY MLGPMTPFLAF 1242

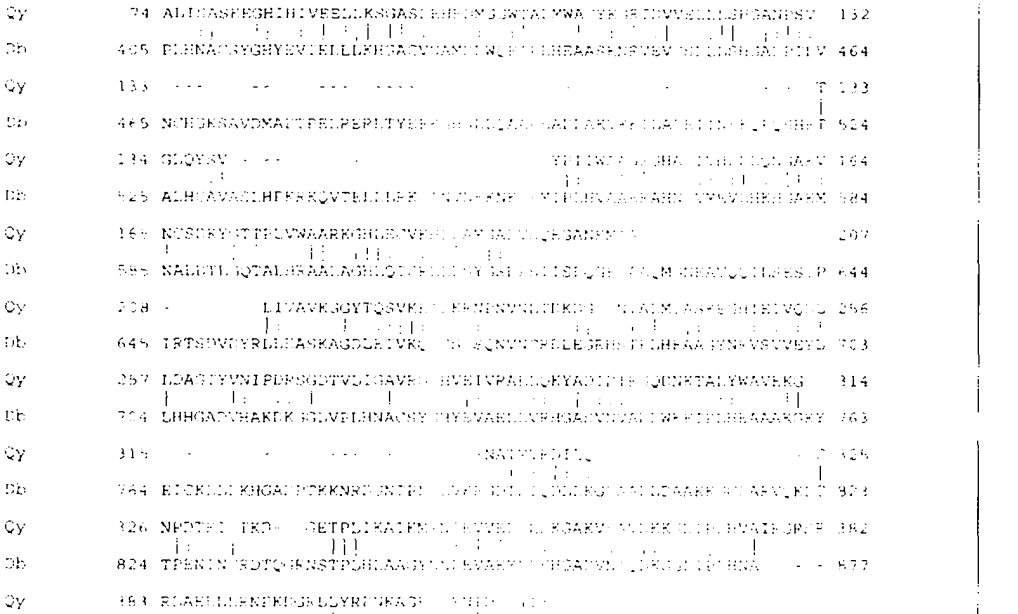
Qy        893  -  --  [illegible]TALN-  [illegible]  -RRDTYRR 907
Db       1243 ERLFFAVFGQTTTLDINPMRHLRPEWTEV[illegible]FKFVPTIYLLVSVVVLINLLIAMMSDTYQR 1302

Qy        908 RQMQR 912
Db       1303 IQMNR 1307
```

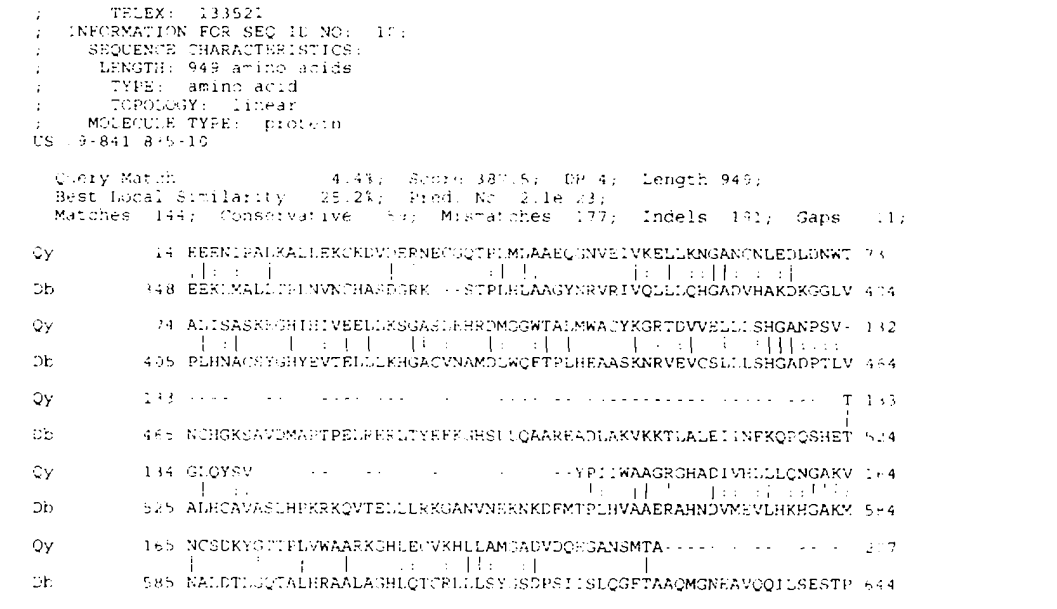
```
RESULT 13
US-09-196-387-10
; Sequence 10, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/[illegible]
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 949 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-196-387-10

  Query Match             4.4%;  Score 387.5;  DB 3;  Length 949;
  Best Local Similarity   25.2%;  Pred. No. 2.1e-23;
  Matches  144;  Conservative   59;  Mismatches  177;  Indels  191;  Gaps   11;

Qy         14 EEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKNGANCNLEDLDNWT 73
Db        348 EEKLMALLTPLNVNCHASDGRK --STPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404
```

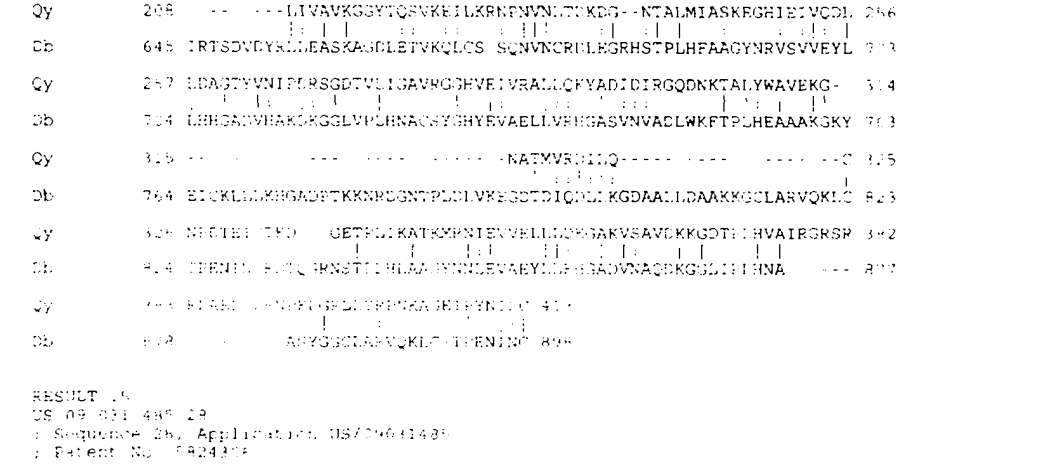


RESULT 14
US 09 841 845-10
; Sequence 10, Application US/09841845
; Patent No. 6506587
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN TH
; TITLE OF INVENTION: OF DNA DAMA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:
; STREET: 41
; CITY: Rockefeller
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07621
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DO
; SOFTWARE: PatentIn Release #1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/194,08
; FILING DATE
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq , David A
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487 5800
; TELEFAX: 212 944-1684



; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 949 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US 09-841 845-10

Query Match 4.4%; Score 38.5; DB 4; Length 949;
Best Local Similarity 25.2%; Pred. No. 2.1e-23;
Matches 144; Conservative 54; Mismatches 177; Indels 191; Gaps 11;



RESULT 15
US 09 831 485 29
; Sequence 29, Application US/09831485
; Patent No. 6924273
; GENERAL INFORMATION:
; APPLICANT: Tabel, Ilana
; APPLICANT: Wilson, B. Scot
; TITLE OF INVENTION: BIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Wind ..  Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/ ...041,...
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08 947,...
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,483
;      REFERENCE/DOCKET NUMBER:  BW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  28
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
08-09-03, 48E 28

  Query Match              4.2%;  Score ...;  Le ..;  D.  ...  ...-,
  Best Local Similarity   34.2%;  Pred. No. ...e-..;
  Matches  113;  Conservative   47;  Mismatches  134;  Indels   ..;  Gaps    4;

Qy          45 AABQ INVEIVKELLKKNIANFNLEIK VWTAFLKAAFNEGTF YAALIEILSLPWRMSS 104
               :    | :: | || :;| : |  : :          :   :     :'     :      |
Db          44 AAPA INLERVLELLRSGTDINTTNKY LNNAIHLAARFGRSEIV-ELIFKAGVNAATRKG 103

Qy         105 WTALMWAVYKGRTDVVELLISRGAN ...  ...YVYEIWAF ...HRA ...LIQN NKV 164
               ||;   :   ::  | : . :|||  .   :     :                 :  :
Db         104 NTALHIASLAGQLEIVTVLVENGAN...  ...PDVYAA ...HHL ...VIAHYAN 160

Qy         165 NCSEKYGTTPLYWAAAKGHLECVKF LAKHAEVNEPDAHGWTPLYVAVQEHKLSVVEV 224
               :  | ::   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         161 ALTRHGHTPLAVA...QGHHEAVVAV...  ...EHGADVNAQ ...EFA ...AAAHGH 218

Qy         225 KRNPVVNLTDKDGNTALMIASKEGH EIVLLLLNAETVVNI ...E ...VNAVKD HVE 284
               :   ' ::  | |     ||:  :  :   |:   :|  :        :
Db         219 QNEHVEIVSKSGFTPLHIAAHYGN ...  ...  ...  ...  ...  ...  ... 278

Qy         285 IVRA ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ... 341
               :
Db         279 MVRI ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ... 335

Qy         342 KMRNIFVVRLLLEKGAEVNAVNYF ...  ...  ... 388
               :  : :|:|  :|:|'  :   :  '  :
Db         336 RSGHIQVVDLLLEKGAPISAKTKN ...  ...  ... 382


Search completed: November 21, 2003, 1...
Job time : 28 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 12:..:04 ; Search time 44.5 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8884
Sequence: 1 MSVDISQSVINYVFRENITA[illegible]PDPEEP[illegible]SIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726577 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

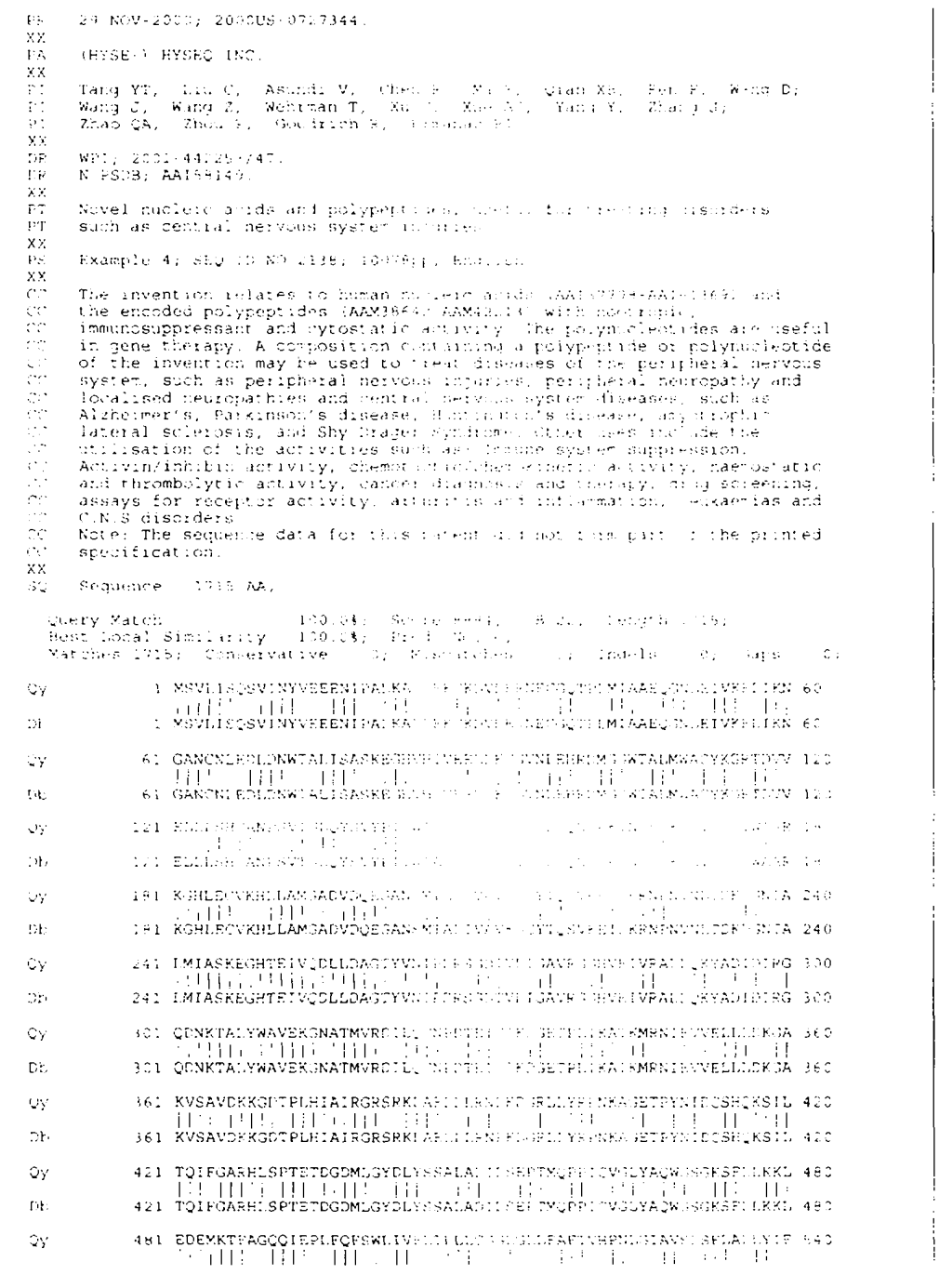
Result No.	Score	Query Match	Length	DB	ID	Description
1	8884	100.0	1715	22	AAM38993	Human polypeptide
2	8884	100.0	1715	22	AAM38995	Human polypeptide
3	8884	100.0	1715	23	AAE25144	Human ARMS protein
4	8884	100.0	1715	23	AAU96448	Human kidins220 pr
5	8884	100.0	1715	24	AAE33128	Human cytoskeleton
6	8362.5	94.1	1762	23	AAU96441	Rat kidins220 prot
7	8301	93.4	1763	23	AAU96244	Rat kidins220 prot
8	8266	93.0	1715	23	AAE25143	Rat ARMS protein.
9	3950.5	44.5	735	22	AAG70604	Human colon cancer
10	2851.5	32.1	551	22	AAE01035	Human death domain
11	2791	31.4	543	22	AAB95191	Human protein sequ
12	2638	29.7	513	22	AAE03645	Human extracellula
13	2603	29.2	624	22	ABG08697	Novel human diagno
14	2379.5	26.8	1438	23	ABB64857	Drosophila melanog
15	2379.5	26.8	1498	23	AAE25146	Fruit fly ARMS pro
16	1832.5	20.6	1398	23	AAE25145	Worm ARMS protein.
17	666	7.5	129	22	AAB94785	Human protein sequ
18	620	7.0	120	22	AAM40811	Human polypeptide
19	553	6.2	109	21	AAB56277	Human secreted pro
20	547.5	6.2	986	22	ABG01785	Human ankyrin 4 po
21	543.5	6.1	1250	24	AAE31667	Human structural a
22	526	5.9	1802	22	AAM79160	Human protein SEQ
23	521	5.9	4274	22	ABG00372	Novel human diagno
24	519.5	5.8	2443	22	ABB60521	Drosophila melanog
25	518.5	5.8	4386	22	ABG07376	Novel human diagno
26	519	5.8	1724	23	AAC20513	Protein of APP rel
27	514	5.8	4397	22	ABG21944	Novel human diagno
28	510.5	5.7	1536	24	ABU11523	Human MDDT polypep
29	499.5	5.6	187	22	AAM40779	Human polypeptide
30	494	5.6	456	22	ABU53033	Cell structure and
31	482.5	5.4	784	22	ABB62206	Drosophila melanog
32	480.5	5.4	1745	19	AAW70608	Full length ankyri
33	480.5	5.4	1745	19	AAW76776	D. immitis ankyrin
34	480.5	5.4	1745	21	AAB11589	D. immitis ankyrin
35	480.5	5.4	1745	23	AAO21368	Ankyrin protein se
36	478.5	5.4	1549	22	ABB58328	Drosophila melanog
37	478.5	5.4	1549	22	ABB67412	Drosophila melanog
38	463	5.2	447	22	ABU53031	Cell structure and
39	463	5.2	489	22	ABU53032	Cell structure and
40	463	5.2	2119	22	ABB64823	Drosophila melanog
41	453.5	5.1	447	22	ABU53030	Cell structure and
42	453	5.1	1056	24	AAE33684	Human structural a
43	447	5.0	743	23	ABB97453	Novel human protei
44	446	5.0	1133	22	ABG28071	Novel human diagno
45	440	5.0	378	22	AAB64963	Human secreted pro

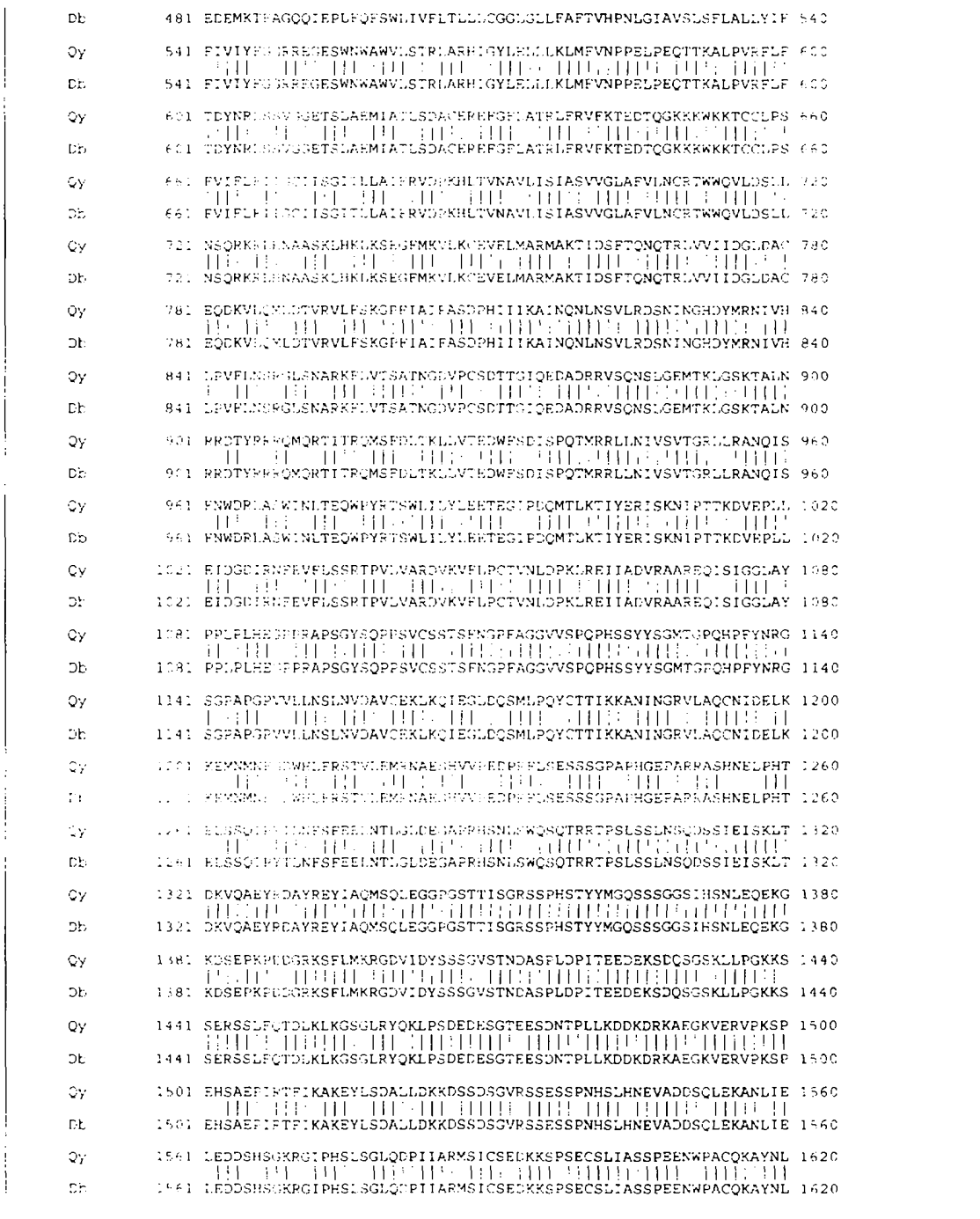
ALIGNMENTS

RESULT 1
AAM38993
ID AAM38993 standard; Protein; 1715 AA.
XX
AC AAM38993;
XX
DT 26-JUL-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2138.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 20-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0553217.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0633450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.

PD 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen F, Ni J, Qian XD, Fan P, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue A, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Steelman P;
XX
DR WPI; 2001-441259/47.
DR N-PSDB; AAI99149.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries.
XX
PS Example 4; SEQ ID NO 3195; 1047pp; English.
XX
CC The invention relates to human nucleic acids (AAI99144-AAI99169) and
CC the encoded polypeptides (AAM38647-AAM38670) with neurogenic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy Drager syndrome. Other uses include the
CC utilisation of the activities such as immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders
CC Note: The sequence data for this patent did not appear in the printed
CC specification.
XX
SQ Sequence 1716 AA;

Query Match 100.0%; Score ####; DB 2; Length 1716;
Best Local Similarity 100.0%; Pred. No. 0.;
Matches 1716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```
Qy       1621 NRTPSTVTLNNNSAPANRANQNPIAMEIPRETIQVTLAPSSFPNPTTIQPENLKSMTHPP 1680
             ||||  ||||||   ||||   |  |  |  |  |  |  |||  |||  ||   |
Db       1621 NRTPSTVTLNNNSAPANRANQNPIAMEIPRETIQVTLAPSSFPNPTTIQPENLKSMIHPP 1680

Qy       1681 SQRSSYTRLSKDFFELHAAASNE 1703
              ||| ||||||| ||||||| ||||   |||
Db       1681 SQRSSYTRLSKDFFELHAAASNE 1703


RESULT 2
AAM39025
ID   AAM39025 standard; Protein; 1715 AA.
XX
AC   AAM39025;
XX
DT   28-OCT-2001 (first entry)
XX
DE   Human polypeptide SEQ ID NO 2170.
XX
KW   Human, nootropic, immunosuppressant, cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy, central nervous system; CNS;
KW   Alzheimer's, Parkinson's disease, Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy Drager syndrome; thrombolytic;
KW   chemokinetic; thrombolytic; anti-coagulant; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2001WO-US49363.
XX
PR   21-JAN-2000; 2000US-0488925.
PR   28-APR-2000; 2000US-0560317.
PR   03-JUL-2000; 2000US-0609042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0633450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   20-NOV-2000; 2000US-0717344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Adams M, Chen Y, Li A, Wang W,
PI   Wang C, Wang Z, Morrison P, Xu
PI   Zhao QA, Zhou P, Ni J,
XX
DR   WPI; 2001-441155/47.
DR   N-PSDB; AAP06181.
XX
PT   Novel nucleic acids and polypeptides useful to treat and diagnose
PT   such as central nervous system disorders
XX
PS   Example 4; SEQ ID NO 2170; 1003pp; English.
XX
CC   The invention relates to human nucleic acids AAP06180-AAP06181 and
CC   the encoded polypeptides (AAM39864-AAM39865). The polypeptides
CC   immunosuppressant and cytostatic activity. The polypeptides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous system neuropathy, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, Amyotrophic
CC   lateral sclerosis, and Shy Drager syndrome. Other uses include the
CC   utilisation of the activities such as immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification.
XX
SQ   Sequence  1715 AA;

  Query Match             100.0%;  Score 8884;  DB 22;  Length 1715;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1715;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSVLISQLPFINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNIEIVKELIKN 60
              ||  |||   ||  |||  |||  ||| |  |||  ||||  |||| ||||  |||||
Db          1 MSVLISQLPFINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNIEIVKELIKN 60

Qy         61 GANCNLEDLDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120
              |||  ||  |||  ||||  |||  |||  ||||||||||  |||||  ||||
Db         61 GANCNLEDLDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
              |||  ||  |||  ||||  |||| |||| |||| |||| |||| |||| |||| |||||
Db        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

Qy        181 KGHLEVVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
              |  |||   ||||||||||  |||| ||||||||||||||||||||||||||||||||||
Db        181 KGHLEVVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy        241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
              |||  ||||||  |||  |||  |||  |||| |||| |||| |||| |||| |||| ||||
Db        241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQCNPDTPICTKDGETPLIKATKMRNIEVVELLLDKGA 360
              |||  |||   |||  |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        301 QDNKTALYWAVEKGNATMVRDILQCNPDTPICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy        361 KVSAVDRHGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
              |||||  |||  |||  |||  |||| |||| |||| |||| |||| |||| ||||||
Db        361 KVSAVDRHGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy        421 TQIPGARHLSPTETDMDMLGYDLYRSALADILSEPTMQPPICVGLYAQWDSGKSFLLKKL 480
              |||  |||  |||  |||  |||| |||| |||| |||| |||| |||| |||| ||||||
Db        421 TQIPGARHLSPTETDMDMLGYDLYRSALADILSEPTMQPPICVGLYAQWDSGKSFLLKKL 480

Qy        481 EDEMKTPATQQIEPLFQFSWLIVPLTLLVGGLGLLPAFTVHPNLGIAVSLSPLALLYIF 540
              |||  ||  |||  |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        481 EDEMKTPAIQQIEPLFQFSWLIVPLTLLVGGLGLLPAFTVHPNLGIAVSLSPLALLYIF 540

Qy        541 PIVIYPRWPPVGESWNWAWVLSTPLASHIDYLELLLKLMPVNPPFLPEQTTKALPVPFLP 600
              |||  |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        541 PIVIYPRWPPVGESWNWAWVLSTPLASHIDYLELLLKLMPVNPPFLPEQTTKALPVPFLP 600

Qy        601 TIYNRLAVPGIETSLAEMIAISGAVFPPFGLATRLFPVEKTEDTQGKKFWKKTCCLPS 660
              |||  |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        601 TIYNRLAVPGIETSLAEMIAISGAVFPPFGLATRLFPVEKTEDTQGKKFWKKTCCLPS 660

Qy        661 PVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720
              |  |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        661 PVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

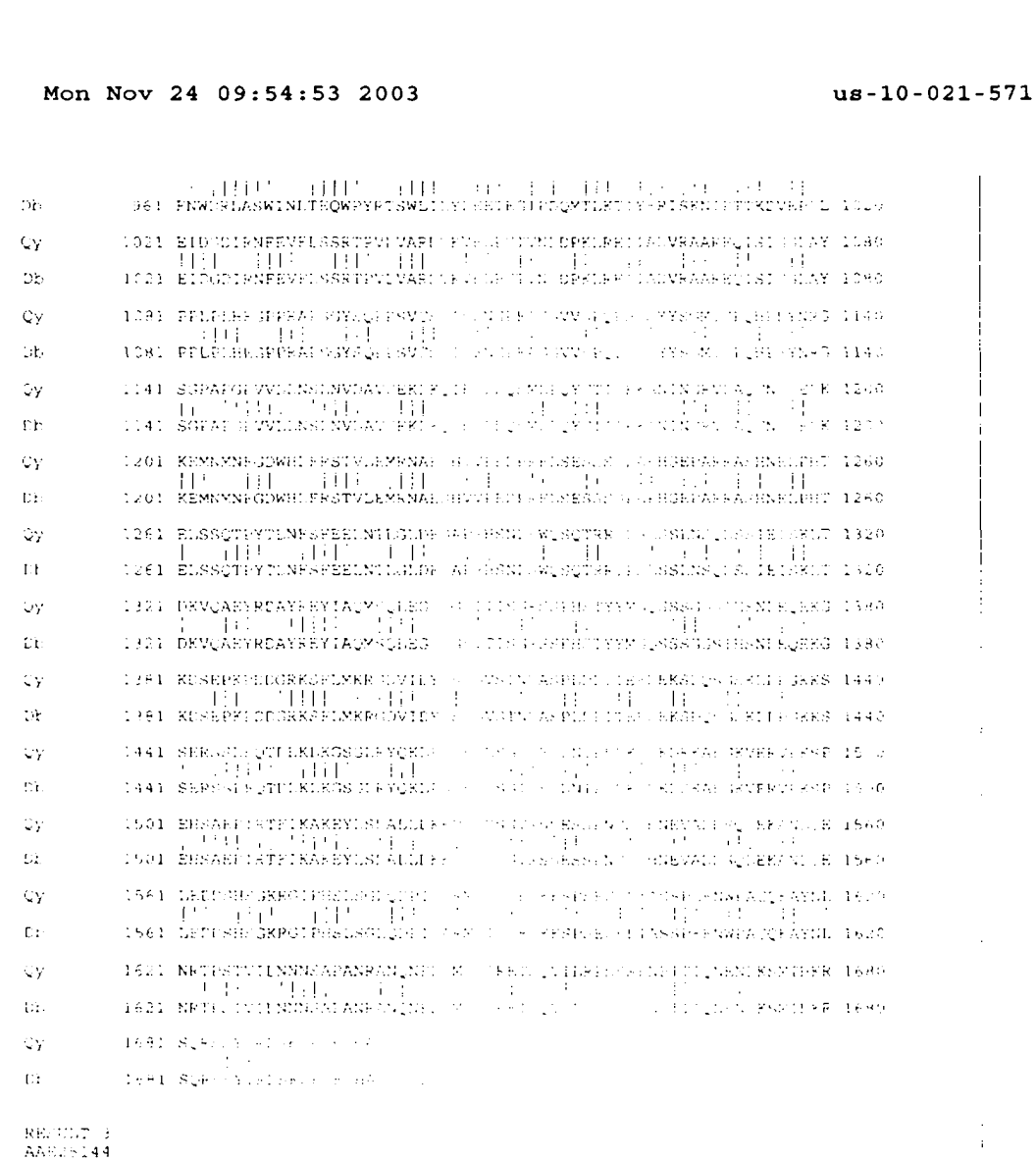
Qy        721 NSQRKPTVNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIDGLDAC 780
              |||  |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        721 NSQRKPTVNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIDGLDAC 780

Qy        781 EQDKVLQVLDTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        781 EQDKVLQVLDTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

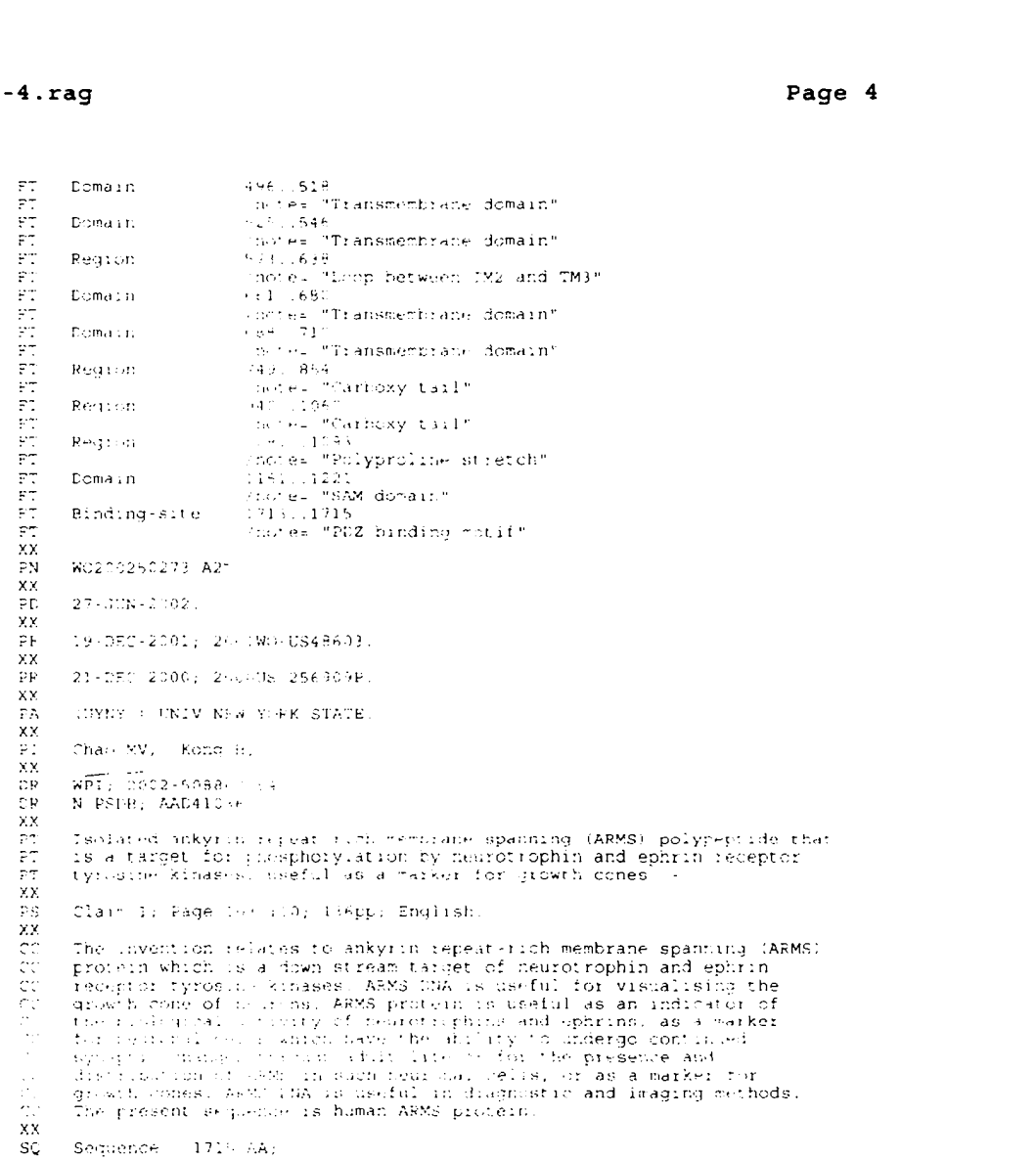
Qy        841 LPVFLNSSGLSNARKFLVTSATNGDVPCSDTTGICEDADRRVSQNSLGEMTKLGSKTALN 900
              |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        841 LPVFLNSSGLSNARKFLVTSATNGDVPCSDTTGICEDADRRVSQNSLGEMTKLGSKTALN 900

Qy        901 RRDTYRPPQMQRTITRQMSPDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960
              |||  |||  |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||||
Db        901 RRDTYRPPQMQRTITRQMSPDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960

Qy        961 PNWLPIAGIIFLEQWEIPTLWLILYLEKEGIPDQMTLKPIYEPISKNIPTTKDVEPLL 1020
```

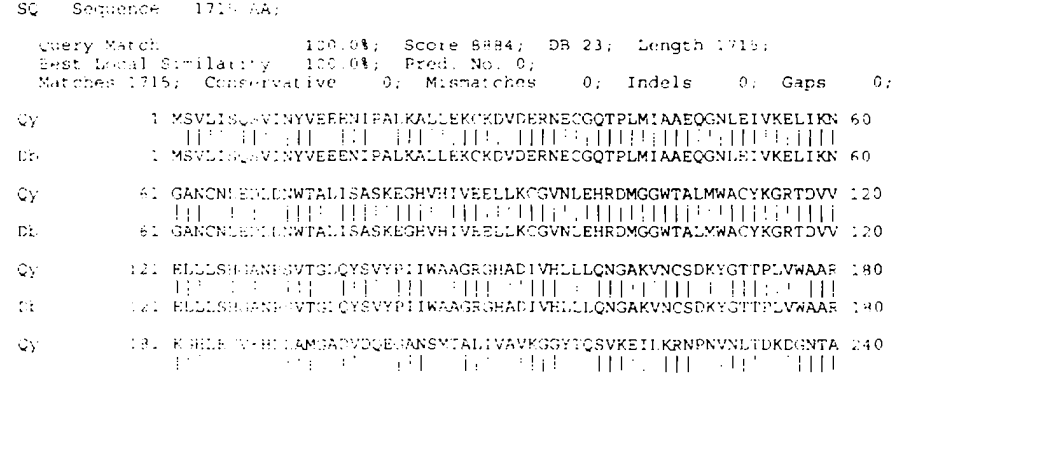


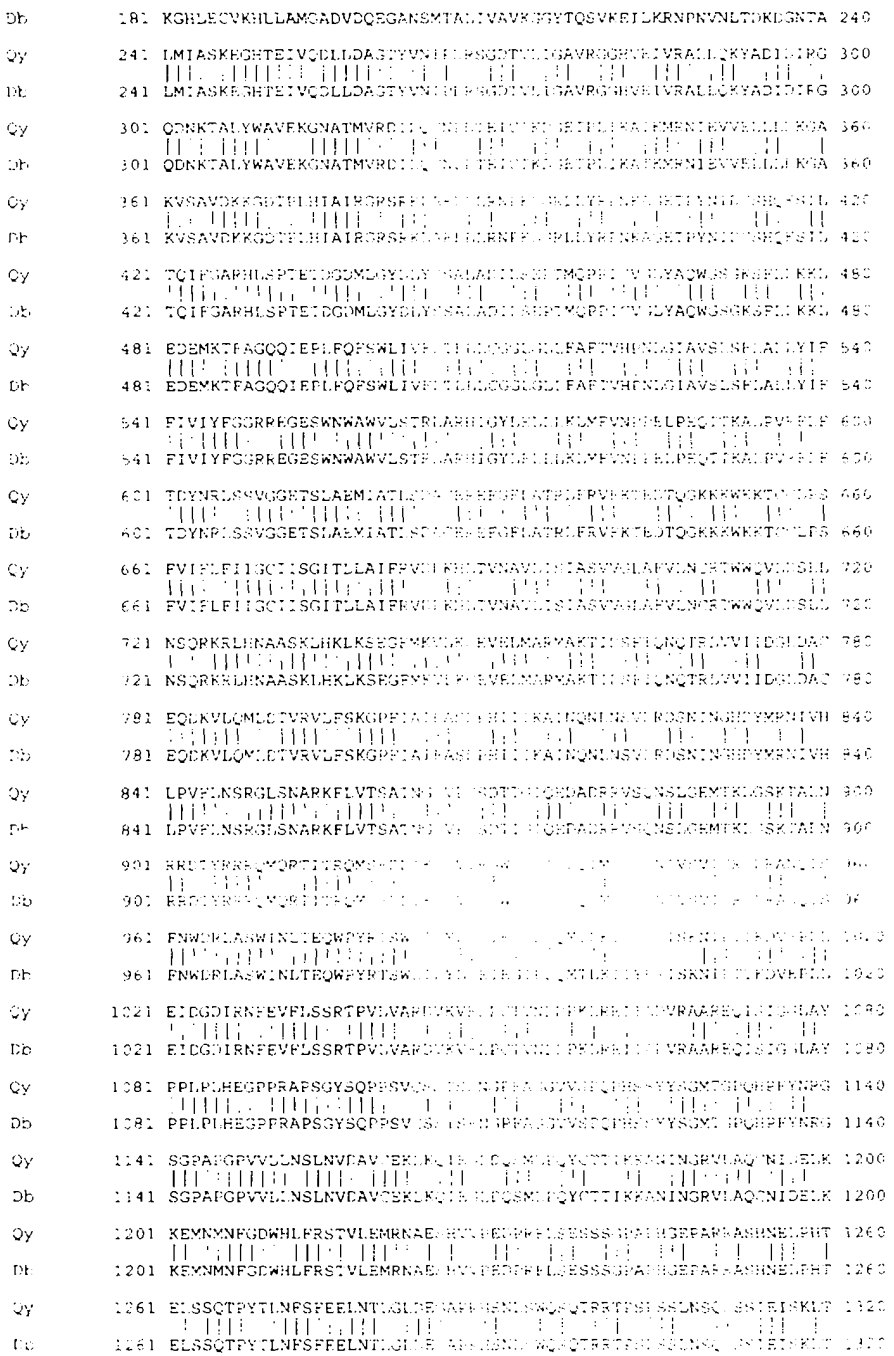
RESULT 3
AAE87144
ID AAE87144 standard; Protein; 1715 AA.
XX
AC AAE87144;
XX
DT 30-OCT-2002 (first entry)
XX
DE Human ARMS protein.
XX
KW Ankyrin repeat rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; [illegible]; cell motility;
KW diagnostic; imaging; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 17..193
FT /note= "Ankyrin repeat"
FT Region 294..493
FT /note= "Ankyrin repeat"

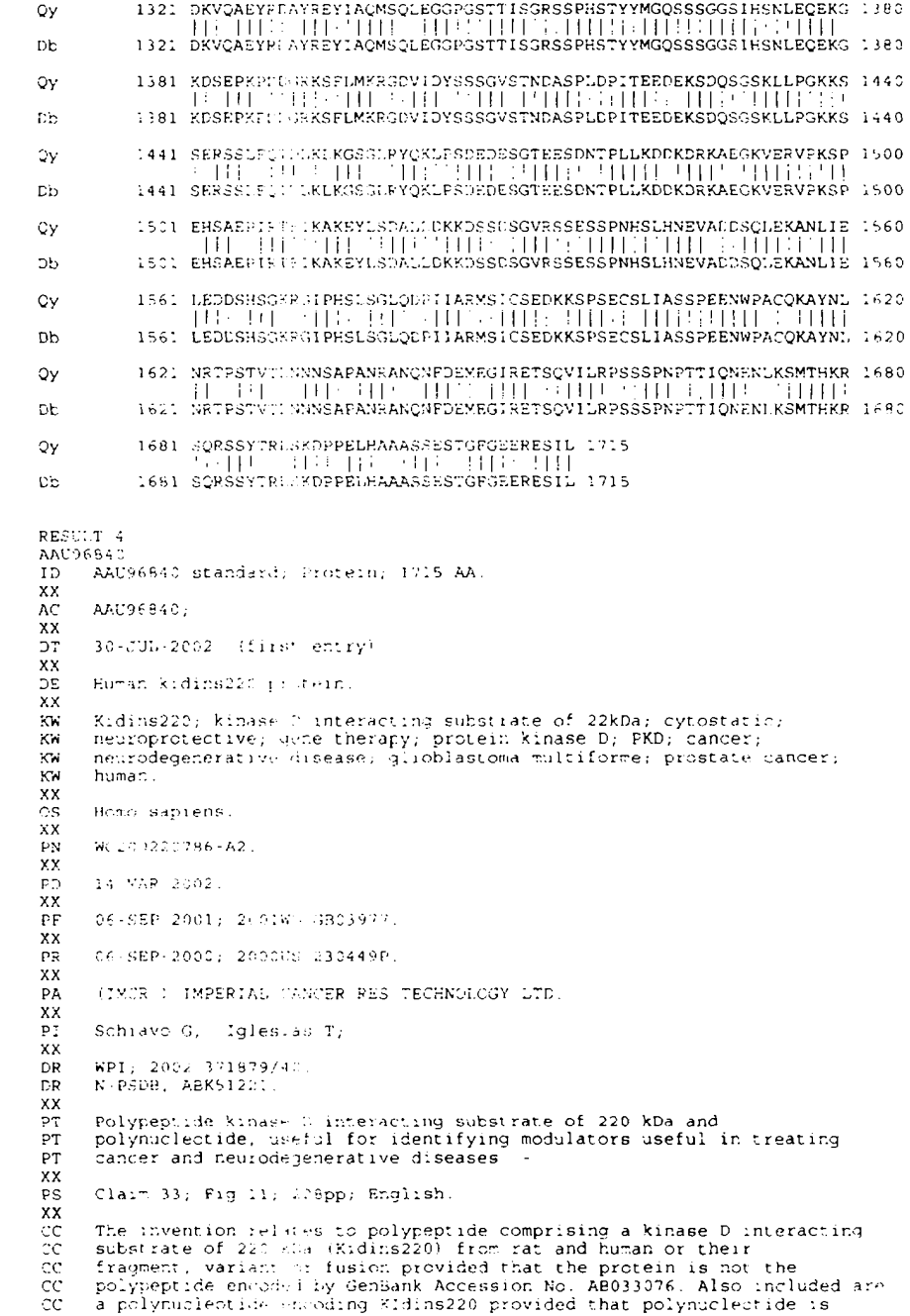


FT Domain 496..518
FT /note= "Transmembrane domain"
FT Domain 546..568
FT /note= "Transmembrane domain"
FT Region 571..638
FT /note= "Loop between TM2 and TM3"
FT Domain 639..660
FT /note= "Transmembrane domain"
FT Domain 694..716
FT /note= "Transmembrane domain"
FT Region 743..854
FT /note= "Carboxy tail"
FT Region 947..1096
FT /note= "Carboxy tail"
FT Region 1..1715
FT /note= "Polyproline stretch"
FT Domain 1151..1221
FT /note= "SAM domain"
FT Binding-site 1711..1715
FT /note= "PDZ binding motif"
XX
PN WO200250273 A2
XX
PD 27-JUN-2002.
XX
PF 19-DEC-2001; 2001WO-US48603.
XX
PR 21-DEC-2000; 2000US-256309P.
XX
PA (UYRY) UNIV NEW YORK STATE.
XX
PI Chan AY, Kong H.
XX
DR WPI; 2002-508804/54.
DR N-PSDB; AAD41054.
XX
PT Isolated ankyrin repeat rich membrane spanning (ARMS) polypeptide that
PT is a target for phosphorylation by neurotrophin and ephrin receptor
PT tyrosine kinases useful as a marker for growth cones -
XX
PS Claim 1; Page 100; 106pp; English.
XX
CC The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
CC protein which is a down stream target of neurotrophin and ephrin
CC receptor tyrosine kinases. ARMS DNA is useful for visualising the
CC growth cone of neurons. ARMS protein is useful as an indicator of
CC the activation activity of neurotrophins and ephrins, as a marker
CC for [illegible] which have the ability to undergo controlled
CC [illegible] and for the presence and
CC distribution of ARMS in such neurons, cells, or as a marker for
CC growth cones. ARMS DNA is useful in diagnostic and imaging methods.
CC The present sequence is human ARMS protein.
XX
SQ Sequence 1715 AA;

Query Match 100.0%; Score 8884; DB 23; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







RESULT 4
AAU96840
ID AAU96840 standard; Protein; 1715 AA.
XX
AC AAU96840;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human Kidins220 protein.
XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer;
KW human.
XX
OS Homo sapiens.
XX
PN WO200225296-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000GB-0330449P.
XX
PA (IMOR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Schiavo G, Iglesias T;
XX
DR WPI; 2002-371879/40.
DR N-PSDB; ABK51201.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating
PT cancer and neurodegenerative diseases -
XX
PS Claim 33; Fig 11; 178pp; English.
XX
CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by GenBank Accession No. AB033076. Also included are
CC a polynucleotide encoding Kidins220 provided that polynucleotide is

```
CC   not any one of the clones corresponding to the 61 GenBank Accession Nos.
CC   given in the specification; an expression/replicable vector comprising
CC   the polynucleotide, a recombinant host cell containing the polynucleotide
CC   or vector; an anti-Kidins220 antibody used in the preparation of
CC   Kidins220); an agent (A1) which modulates activity of Kidins220,
CC   protein kinase D (PKD) or interaction between PKD and Kidins220;
CC   a mutant animal transgenic for kidins220, the use of an agent capable of
CC   detecting the expression of Kidins220 gene in manufacture of a diagnostic
CC   reagent for diagnosing or prognosing cancer or for monitoring the
CC   progression of cancer in a patient; a compound comprising a moiety that
CC   selectively binds to Kidins220 or its variant and another moiety; and a
CC   nucleic acid molecule encoding the compound. The antibody is useful for
CC   modulating the activity of Kidins220, where the antibody prevents a
CC   region of Kidins220 interacting from another protein, or reduces the
CC   ability of Kidins220 to bind to adenosine triphosphate (ATP). The
CC   kidins220 polynucleotide is useful for identifying an agent which
CC   modulates the promoter activity of the polynucleotide, and Kidins220 is
CC   useful for identifying an agent which modulates activity or Kidins220,
CC   protein kinase D (PKD) and the interaction between PKD with Kidins220.
CC   Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
CC   therapy or for diagnosing or monitoring progression of cancer in a
CC   patient. The agent is useful in manufacture of medicament, for use in
CC   treatment of neurodegenerative disease, and the agent which inhibits
CC   function of Kidins220 gene or its product is useful for treating cancer
CC   which is glioblastoma multiforme or prostate cancer. The compound is
CC   useful for imaging cancer in an ex vivo, and for diagnosing or
CC   prognosing, and also for treating an individual. The present
CC   sequence represents Human kidins.."
XX
SQ   Sequence  1715 AA,

  Query Match             100.0%;  Score 9864;  DB 2;  Length 1715;
  Best Local Similarity   100.0%;  Pred. No. 0.;
  Matches 1715;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSVLISQSVINYVEENIFALFA...  60
              ||  |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db          1 MSVLISQSVINYVEENIFALFA...  60

Qy         61 GANSNLEDLFNWTALISASKEP...  120
              ||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db         61 GANSNLEDLFNWTALISASKEP...  120

Qy        121 ELLLSHGANPSVTGLQYSVYPF...  180
              ||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        121 ELLLSHGANPSVTGLQYSVYPF...  180

Qy        181 KSIFEVFELAMVEGCVELVKYL...  240
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        181 KSIFEVFELAMVEGCVELVKYL...  240

Qy        241 LMIASKEGHTEIVQDLLDAGTPV...  300
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        241 LMIASKEGHTEIVQDLLDAGTPV...  300

Qy        301 QDNKTALYWAVEKGNATMVRDI...  360
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        301 QDNKTALYWAVEKGNATMVRDI...  360

Qy        361 KVSAVDKKGDTPLHIAIRGRSH...  420
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        361 KVSAVDKKGDTPLHIAIRGRSH...  420

Qy        421 TQIFGARHLSPTETDGDMLGYD...  480
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        421 TQIFGARHLSPTETDGDMLGYD...  480

Qy        481 EDEMKTFAGQQIEPLFQFSWLI...  540
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        481 EDEMKTFAGQQIEPLFQFSWLI...  540

Qy        541 FIVIYFGGRREGESWNWAWVLS...  600
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        541 FIVIYFGGRREGESWNWAWVLS...  600

Qy        601 TDYNRLSFVGGETSLAEMIATL...  660
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        601 TDYNRLSFVGGETSLAEMIATL...  660

Qy        661 FVIFLFIIFFISGIILAIFRVD...  720
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        661 FVIFLFIIFFISGIILAIFRVD...  720

Qy        721 NSQRKRLFQAASKLHKLKSSGF...  780
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        721 NSQRKRLFQAASKLHKLKSSGF...  780

Qy        781 EQDKVLQMLPTVRVLFSKGPFI...  840
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        781 EQDKVLQMLPTVRVLFSKGPFI...  840

Qy        841 LPVFLNSPQLSNARKFLVTSAT...  900
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        841 LPVFLNSPQLSNARKFLVTSAT...  900

Qy        901 RRDTYRRPLMQRTITRQMSFDL...  960
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        901 RRDTYRRPLMQRTITRQMSFDL...  960

Qy        961 FNWDRLASWINLTEQWPYRTSW...  1020
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db        961 FNWDRLASWINLTEQWPYRTSW...  1020

Qy       1021 EIDGDIRNPEVPLSSRTPVLVA...  1080
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1021 EIDGDIRNPEVPLSSRTPVLVA...  1080

Qy       1081 PPLPLHEQPPAPSGYSQPPSVC...  1140
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1081 PPLPLHEQPPAPSGYSQPPSVC...  1140

Qy       1141 SGPAPGPVVLLNSLNVDAVCEK...  1200
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1141 SGPAPGPVVLLNSLNVDAVCEK...  1200

Qy       1201 KEMNMNPFQWHLPRSTVLEMRN...  1260
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1201 KEMNMNPFQWHLPRSTVLEMRN...  1260

Qy       1261 EISSQTPFQEEINTALLDEIAP...  1320
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1261 EISSQTPFQEEINTALLDEIAP...  1320

Qy       1321 DKVQAEYPAYREYIAQMSQLEG...  1380
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1321 DKVQAEYPAYREYIAQMSQLEG...  1380

Qy       1381 KDSEPKPFQWRKSFLMKRGDVI...  1440
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1381 KDSEPKPFQWRKSFLMKRGDVI...  1440

Qy       1441 SERSSLPQTPQLKLKGSGLRYQ...  1500
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1441 SERSSLPQTPQLKLKGSGLRYQ...  1500

Qy       1501 EHSAEPIRTPIKAKEYLSDALL...  1560
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1501 EHSAEPIRTPIKAKEYLSDALL...  1560

Qy       1561 LEDDSHSGRPGIPHSLSGLQDP...  1620
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1561 LEDDSHSGRPGIPHSLSGLQDP...  1620

Qy       1621 NRTPSTVTPNNNSAPANRANQNF...  1680
              |||| |||| |||| |||| |||| ||| |||| |||| |||| |||| |||| ||||
Db       1621 NRTPSTVTPNNNSAPANRANQNF...  1680
```

```
Qy       1681 SQRSSYTRLSKDPPELHAAASSE[illegible]SIL 1715
             ||||| . ||||||' |||| .||| ..|
Db       1681 SQRSSYTRLSKDPPELHAAASPE[illegible]SIL 1715


RESULT 5
AAB32128
ID   AAB32128 standard; Protein; 1715 AA.
XX
AC   AAB32128;
XX
DT   24-MAR 2003  (first entry)
XX
DE   Human cytoskeleton associated protein CSAP 26
XX
KW   Human; cytoskeleton associated protein, CSAP 26, atherosclerosis;
KW   cancer; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          42..96
FT                   /note= "Domain present in A 1"
FT   Domain          354..392
FT                   /note= "Domain present in A 1"
FT   Binding site    467..474
FT                   /note= "ATP/GTP binding site motif A"
FT   Domain          494..514
FT                   /note= "Transmembrane domain"
FT   Domain          524..544
FT                   /note= "Transmembrane domain"
FT   Domain          654..674
FT                   /note= "Transmembrane domain"
FT   Domain          687..707
FT                   /note= "Transmembrane domain"
XX
PN   WO200279404-A2.
XX
PD   10-OCT-2002
XX
PF   25 MAR-2002; 2002WO-US08288.
XX
PR   29-MAR-2001; 2001US-280508P.
PR   03-APR-2001; 2001US-281323P.
PR   13 APR-2001; 2001US-283769P.
PR   04 MAY-2001; 2001US-288[illegible]P.
PR   10-MAY-2001; 2001US-[illegible]P.
PR   18-MAY-2001; 2001US-[illegible]P.
PR   29 MAY-2001; 2001US-[illegible]P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hafalia AJA,  Tang TY,  Yue H,  Khan FA,  [illegible],  Duggan BM,
PI   Warren BA,  Duggan BM,  Thangavelu K,  [illegible],  Azimzai Y;
PI   Elliott VS,  Burford N,  Ding L,  Yue H,  Padma S,  [illegible] BM;
PI   Richardson TW,  Lee SY,  Bandman O,  [illegible],  [illegible] Gietzen KJ,
PI   Walia NK,  Griffin JA,  Lee EA,  Swarnakar A,  Rieger RD,  Jones KA,
XX
DR   WPI; 2003-092894/08.
DR   N-PSDB; AAD49615.
XX
PT   New human cytoskeleton-associated proteins, useful in preparing a
PT   composition for diagnosing or treating a disease or condition
PT   associated with decreased expression or overexpression of functional
PT   CSAP e.g., cancer -
XX
PS   Claim 1; Page 195-193; 233pp; English.
XX
CC   The invention relates to new human cytoskeleton associated protein
CC   (CSAP) and its polynucleotide. The polypeptide is useful for preparing
CC   a composition for diagnosing or treating a disease or condition
CC   associated with decreased expression or overexpression of functional
CC   CSAP e.g. atherosclerosis or cancer. The present sequence is human
CC   CSAP-26 protein. The invention is useful in gene therapy.
XX
SQ   Sequence   1715 AA;

  Query Match           100.0%;  Score 8804;  DB 24;  Length 1715;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1715;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60
              ||.|.|| .|| | |||  |||| |||||||||||||||  |||||.|||||.
Db          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Qy         61 GANCNLELINWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120
              ||| .|| ||  ||| .|||  |||||||||| ||||  ||||||| |
Db         61 GANCNLELINWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
              ||| .|| |  || .|| |||  |||| | |||| | |||| |||||| ||
Db        121 ELLLSHGANISVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

Qy        181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
              ||| ||| |||| ||| |||| .|||| |||| |||| .|||||
Db        181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy        241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
              |||||| ||| ||| |||| |||| ||||||||||||| |||| |||||
Db        241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
              |||||| ||| |||| |||| |.||| | ||||.|||||||||||| | ||||||
Db        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy        361 KVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
              |||||. ||| |||| |||| |.||| | ||||.|||||||||||| | ||||||
Db        361 KVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy        421 TQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480
              |||| |||  |||| |||| |||| |||| |||||||||||||| |||||||| |||
Db        421 TQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480

Qy        481 EDEMKTFAQQIEPLFQFSWLIVPLTLLLCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540
              |. ||| ||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        481 EDEMKTFAQQIEPLFQFSWLIVPLTLLLCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540

Qy        541 FIVIYFPMDEGESWNWAWVLSTPLARHIGYLELLLKLMFVNPFELPEQTTKALPVRFLF 600
              .|| |.. ||| .|| |||| |||| |||| |||| |||| |||| ||||
Db        541 FIVIYFPMDEGESWNWAWVLSTPLARHIGYLELLLKLMFVNPFELPEQTTKALPVRFLF 600

Qy        601 IDYNRLSSVSETSLAEMIATLSDACERFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660
              .|| .|. ||| |.| .|| |||| |||| |||| |||| |||| ||||
Db        601 TDYNRLSSVSETSLAEMIATLSDACERFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660

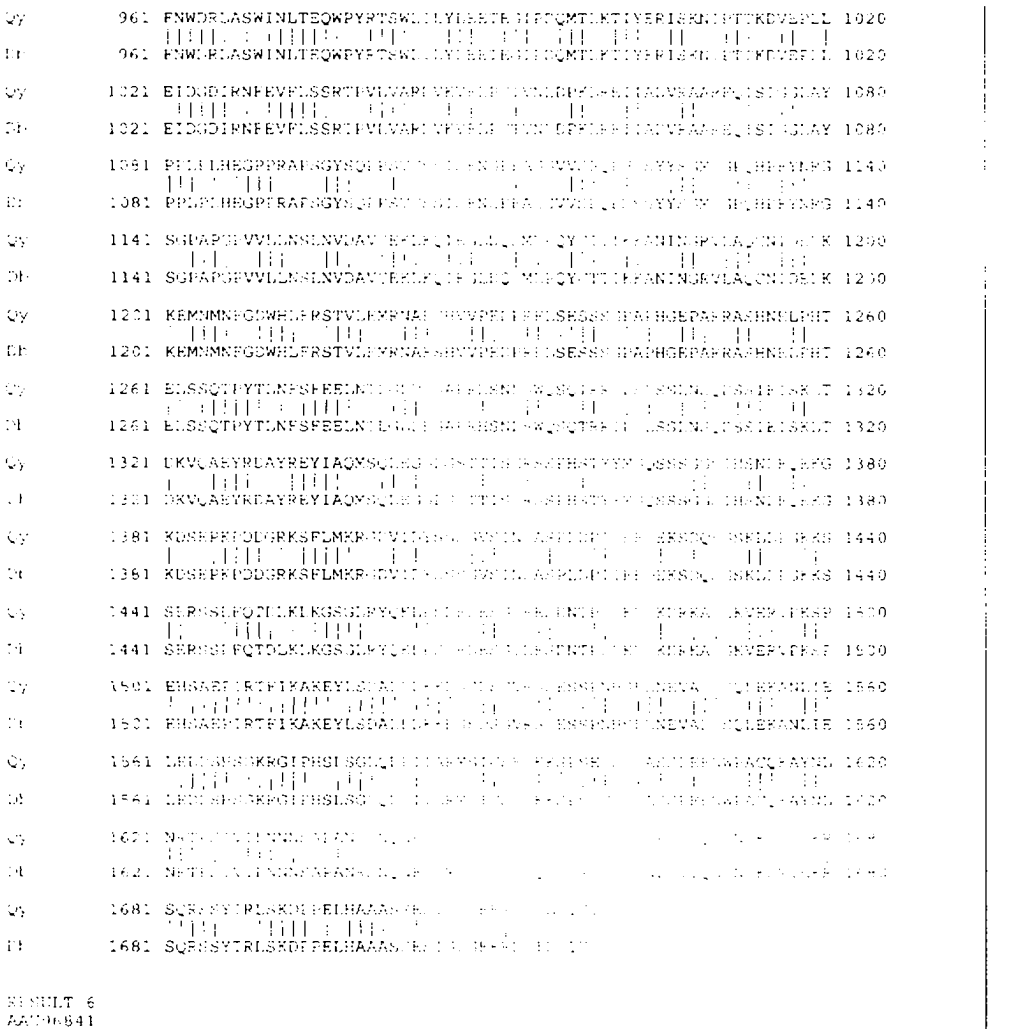
Qy        661 FVIFLFIIGSIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720
              ||. |.| ||| ||| |||| |||| |||| |||| |||| |||| ||||
Db        661 FVIFLFIIGSIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780
              |.|| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        721 NSQRKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy        781 EQDKVLQMLDTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              |||| || |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        781 EQDKVLQMLDTVRVLFSKGPFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Qy        841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900
              |||||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900

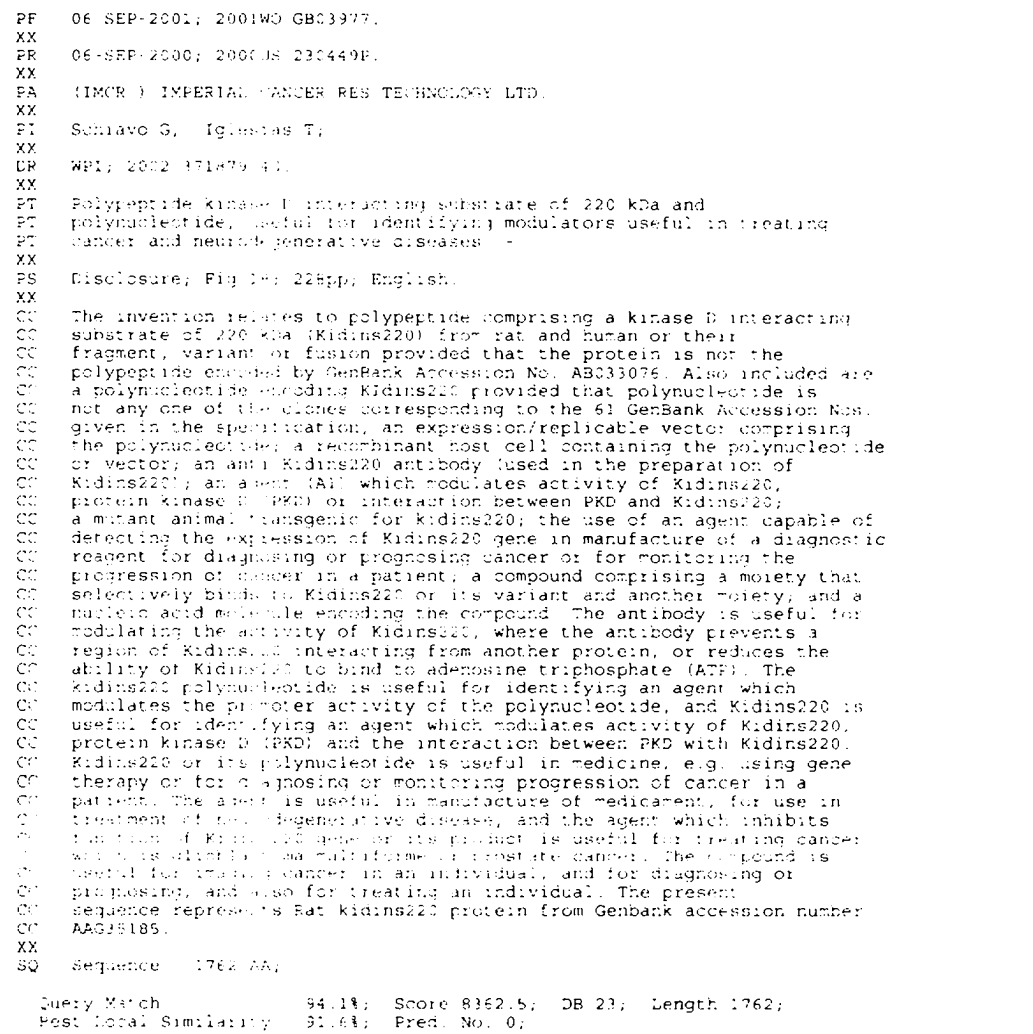
Qy        901 RRDTYRRRQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960
              |.|||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        901 RRDTYRRRQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960
```

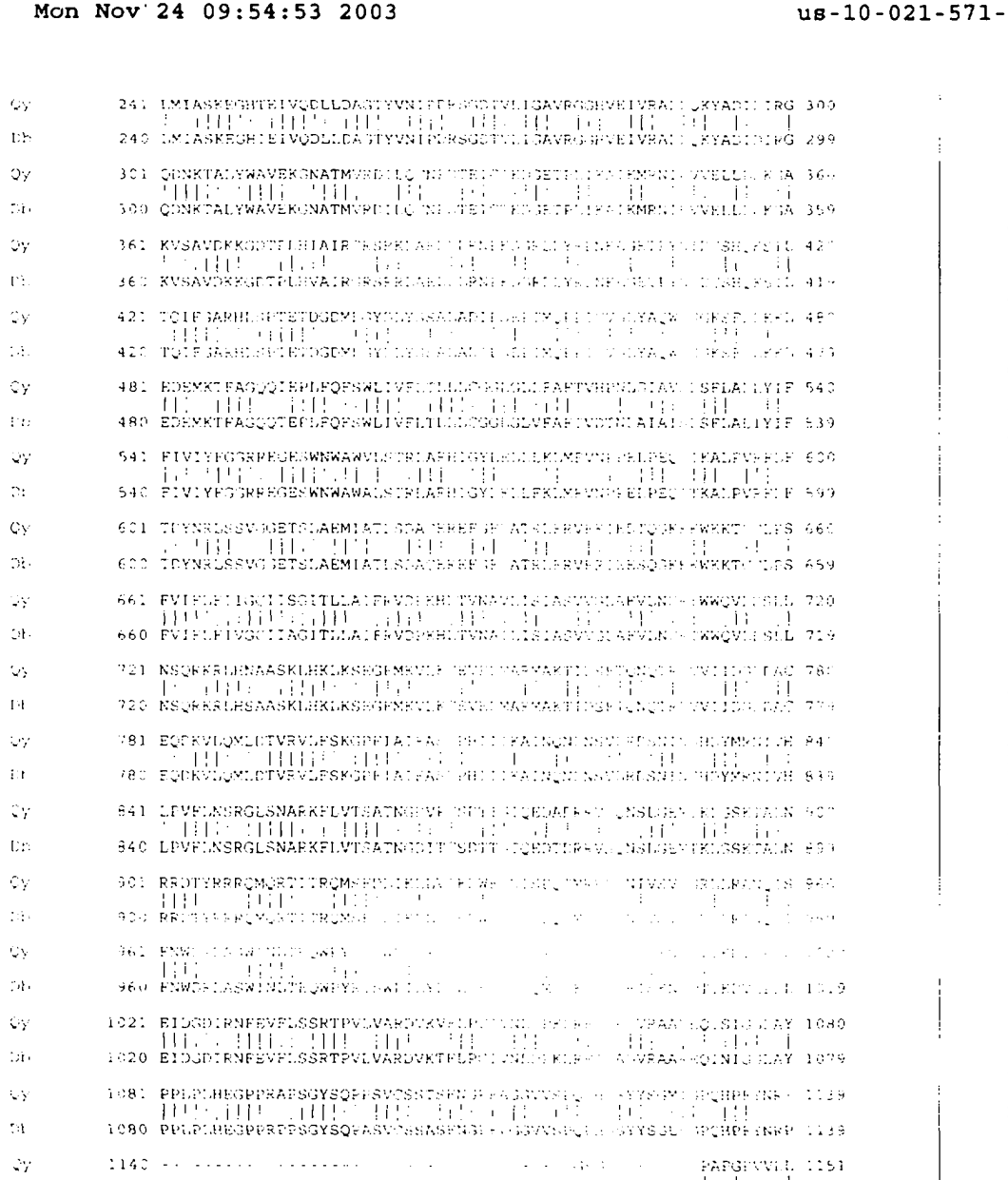


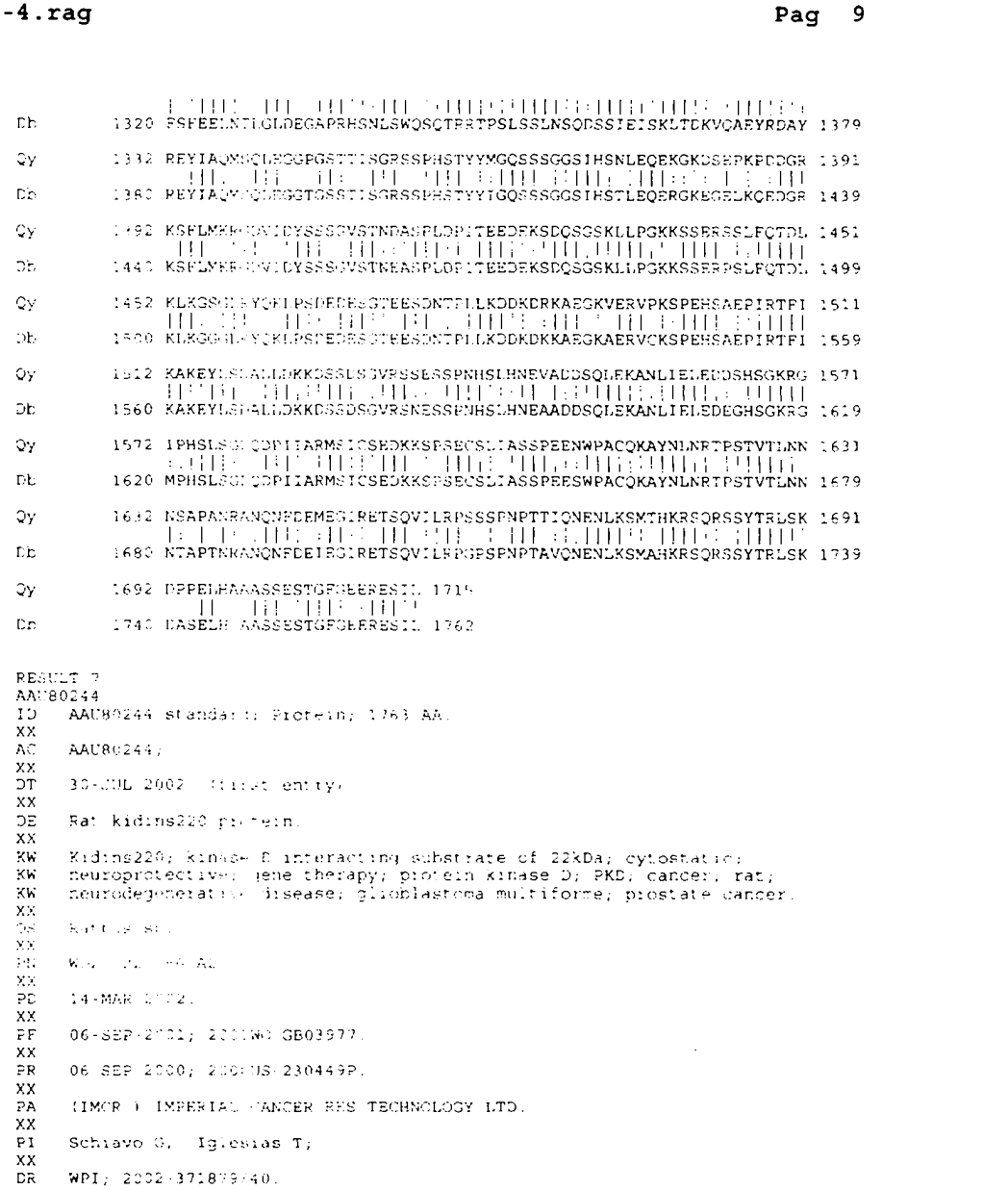
RESULT 6
AAU96841
ID AAU96841 standard; Protein; [illegible] AA.
XX
AC AAU96841;
XX
DT 30-JUL-2002 (first entry)
XX
DE Rat Kidins220 protein from [illegible].
XX
KW Kidins220; kinase D interacting [illegible]; [illegible]
KW neuroprotective; gene therapy; [illegible]
KW neurodegenerative disease; [illegible].
XX
OS Rattus sp.
XX
PN WO200220786-A2.
XX
PD 14-MAR-200[illegible].

PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000US-230449P.
XX
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Scholavo G, Ig[illegible] T;
XX
DR WPI; 2002-371479/41.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating
PT cancer and neurodegenerative diseases -
XX
PS Disclosure; Fig 1+; 228pp; English.
XX
CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by GenBank Accession No. AB033076. Also included are
CC a polynucleotide encoding Kidins220 provided that polynucleotide is
CC not any one of the clones corresponding to the 61 GenBank Accession Nos.
CC given in the specification, an expression/replicable vector comprising
CC the polynucleotide; a recombinant host cell containing the polynucleotide
CC or vector; an anti-Kidins220 antibody used in the preparation of
CC Kidins220; an agent (A1) which modulates activity of Kidins220,
CC protein kinase D (PKD) or interaction between PKD and Kidins220;
CC a mutant animal transgenic for kidins220; the use of an agent capable of
CC detecting the expression of Kidins220 gene in manufacture of a diagnostic
CC reagent for diagnosing or prognosing cancer or for monitoring the
CC progression of cancer in a patient; a compound comprising a moiety that
CC selectively binds to Kidins220 or its variant and another moiety; and a
CC nucleic acid molecule encoding the compound. The antibody is useful for
CC modulating the activity of Kidins220, where the antibody prevents a
CC region of Kidins220 interacting from another protein, or reduces the
CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The
CC kidins220 polynucleotide is useful for identifying an agent which
CC modulates the promoter activity of the polynucleotide, and Kidins220 is
CC useful for identifying an agent which modulates activity of Kidins220,
CC protein kinase D (PKD) and the interaction between PKD with Kidins220.
CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene
CC therapy or for diagnosing or monitoring progression of cancer in a
CC patient. The agent is useful in manufacture of medicament, for use in
CC treatment of neurodegenerative disease, and the agent which inhibits
CC [illegible] of Kidins220 gene or its product is useful for treating cancer
CC [illegible] multiforme, prostate cancer. The compound is
CC useful for treating cancer in an individual, and for diagnosing or
CC prognosing, and also for treating an individual. The present
CC sequence represents Rat kidins220 protein from Genbank accession number
CC AAG25185.
XX
SQ Sequence 1762 AA;

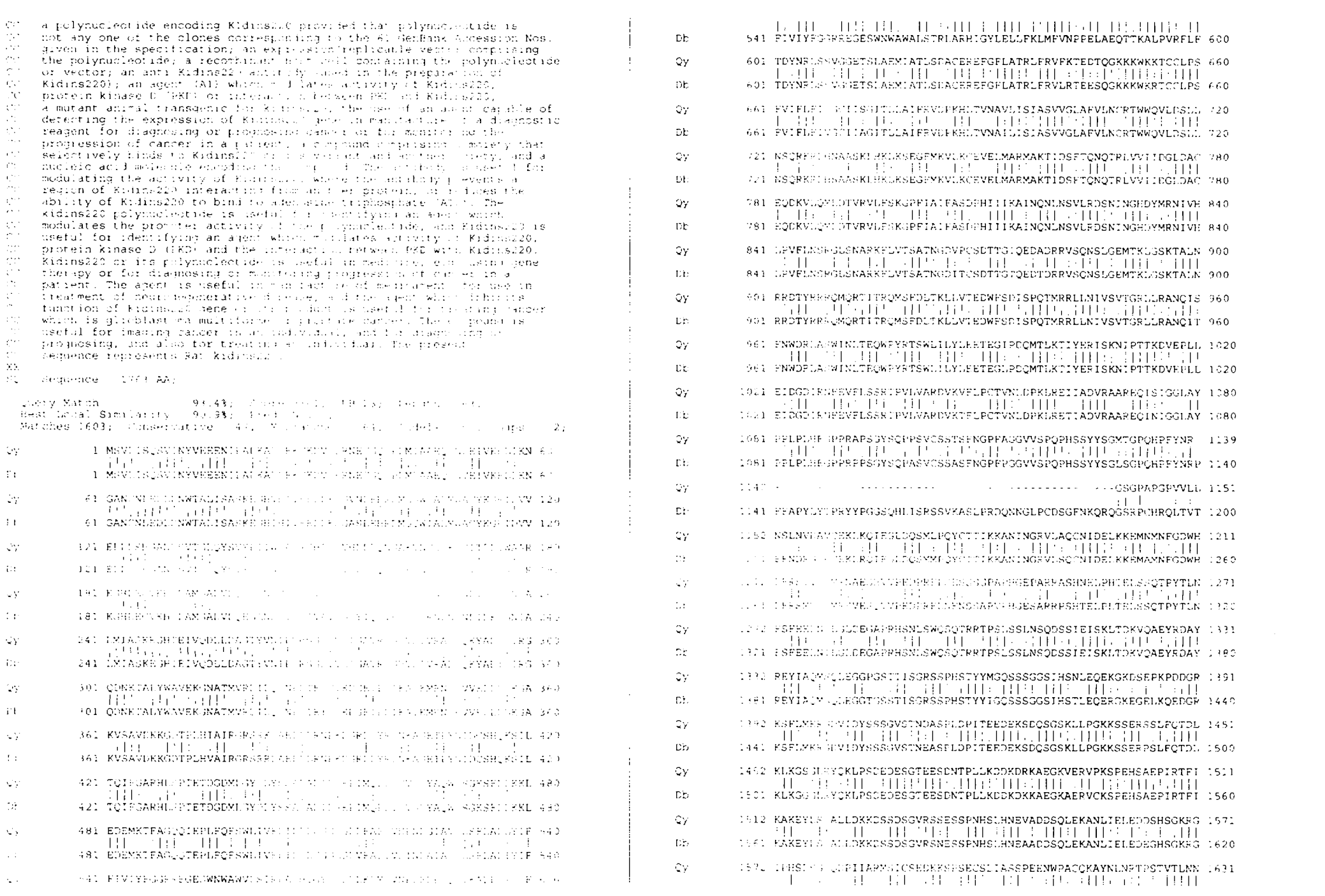
Query Match 94.1%; Score 8362.5; DB 23; Length 1762;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1615; Conservative 47; Mismatches 50; Indels 51; Gaps 4;







RESULT 7
AAU80244
ID AAU80244 standard; Protein; 1763 AA.
XX
AC AAU80244;
XX
DT 30-JUL 2002 (first entry)
XX
DE Rat kidins220 protein.
XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX
OS Rattus sp.
XX
PN WO200271871-A2.
XX
PD 14-MAR 2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000US-230449P.
XX
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Schiavo G, Ig.eias T;
XX
DR WPI; 2002-371879/40.
DR N-PSDB; ABK51214.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating
PT cancer and neurodegenerative diseases -
XX
PS Claim 33; Fig 1; 228pp; English.
XX
CC The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their
CC fragment, variant or fusion provided that the protein is not the
CC polypeptide encoded by GenBank Accession No. AB033076. Also included are



```
Db        1621 MPRSLSGLQDPIIARMSICSEDKRSPSCSLIASSPEESWPAQQKAYNLTRTPSTVTLNN 1680

Qy        1632 NSAPANRANQNPDEMEGIRETSQVILPIGSSFNPTTIQNENIKSMTHKPQRSSYTRLSK 1691

Db        1681 NTAPTNRANQNPDEIEGIRETSQVILPPGPSPNPTAVQNENIKSMAHKPQRSSYTRLSK 1740

Qy        1692 DPPELHAAAASPSTGPGEEPESIL 1715

Db        1741 DASELHAAASPSTGPGEEPESIL 1765


RESULT 8
AAE25143
ID   AAE25143 standard; Protein; 1715 AA.
XX
AC   AAE25143;
XX
DT   30-OCT-2002  (first entry)
XX
DE   Rat ARMS protein.
XX
KW   Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
KW   ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW   diagnostic; imaging; rat.
XX
OS   Rattus norvegicus.
XX
FH   Key             Location/Qualifiers
FT   Region          17..397
FT                   /note= "Ankyrin repeat"
FT   Region          354..493
FT                   /note= "N-terminal region"
FT   Domain          496..519
FT                   /note= "Transmembrane domain"
FT   Domain          525..546
FT                   /note= "Transmembrane domain"
FT   Region          573..618
FT                   /note= "Loop between TM1 and TM2"
FT   Domain          661..680
FT                   /note= "Transmembrane domain"
FT   Domain          688..710
FT                   /note= "Transmembrane domain"
FT   Region          749..864
FT                   /note= "Carboxy tail"
FT   Region          940..1060
FT                   /note= "Carboxy tail"
FT   Region          1091..1135
FT                   /note= "Polyproline region"
FT   Domain          1171..1221
FT                   /note= "SAM domain"
FT   Binding site    1713..1715
FT                   /note= "PDZ binding motif"
XX
PN   WO200250273-A2.
XX
PD   27-JUN-2002.
XX
PF   19-DEC-2001; 2001WO-US48603.
XX
PR   21-DEC-2000; 2000US-256909P.
XX
PA   (UYNY ) UNIV NEW YORK STATE.
XX
PI   Chao MV,  Kong H;
XX
DR   WPI; 2002-508800/54.
DR   N-PSDB; AAD41035.
XX
PT   Isolated ankyrin repeat-rich membrane spanning (ARMS) polypeptide that
PT   is a target for phosphorylation by neurotrophin and ephrin receptor
PT   tyrosine kinases, useful as a marker for growth cones -
XX
PS   Claim 1; Page 87-94; 136pp; English.
XX
CC   The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
CC   protein which is a down stream target of neurotrophin and ephrin
CC   receptor tyrosine kinases. ARMS DNA is useful for visualising the
CC   growth cone of neurons. ARMS protein is useful as an indicator of
CC   the biological activity of neurotrophins and ephrins, as a marker
CC   for neuronal cells which have the ability to undergo continued
CC   synaptic changes through adult life or for the presence and
CC   distribution of ARMS in such neuronal cells, or as a marker for
CC   growth cones. ARMS DNA is useful in diagnostic and imaging methods.
CC   The present sequence is rat ARMS protein.
XX
SQ   Sequence   1715 AA;

  Query Match             93.0%;  Score 8266;  DB 23;  Length 1715;
  Best Local Similarity   92.7%;  Pred. No. 0;
  Matches 1591;  Conservative   53;  Mismatches   70;  Indels    2;  Gaps    2;

Qy          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Db          1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGNVEIVKELLKN 60

Qy         61 GANCNLEDLDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Db         61 GANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACYKGRTDVV 120

Qy        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

Db        121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

Qy        181 KGHLEIVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Db        181 KGHLEIVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy        241 LMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Db        241 LMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Db        301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

Qy        361 KVSAVDKFGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Db        361 KVSAVDKFGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy        421 TQIFGVAPLNSPTETDGMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480

Db        421 TQIFGVAPLNSPTETDGMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480

Qy        481 EDEMKTFAGQQIEPLFQFSWLIVFLTLLLCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540

Db        481 EDEMKTFAGQQTEPLFQFSWLIVFLTLLLCGGLGLVFAFPVDTNLAIAISLSFLALIYIF 540

Qy        541 FIVIYFGSRREGESWNWAWVLSTRLARHIGYLELLLKLMFVNPPELPEQTTKALPVRFLF 600

Db        541 FIVIYFGSRREGESWNWAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFLF 600

Qy        601 TDYNRLSSVGGETSLAEMIATLSDACERFGFLATRLFRVFKTEDTQGKKKWKKTCCLPS 660

Db        601 TDYNRLSSVGGETSLAEMIATLSDACERFGFLATRLFRVFRTEESQGKKKWKKTCCLPS 660

Qy        661 FVIFLFIIGCIIAGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

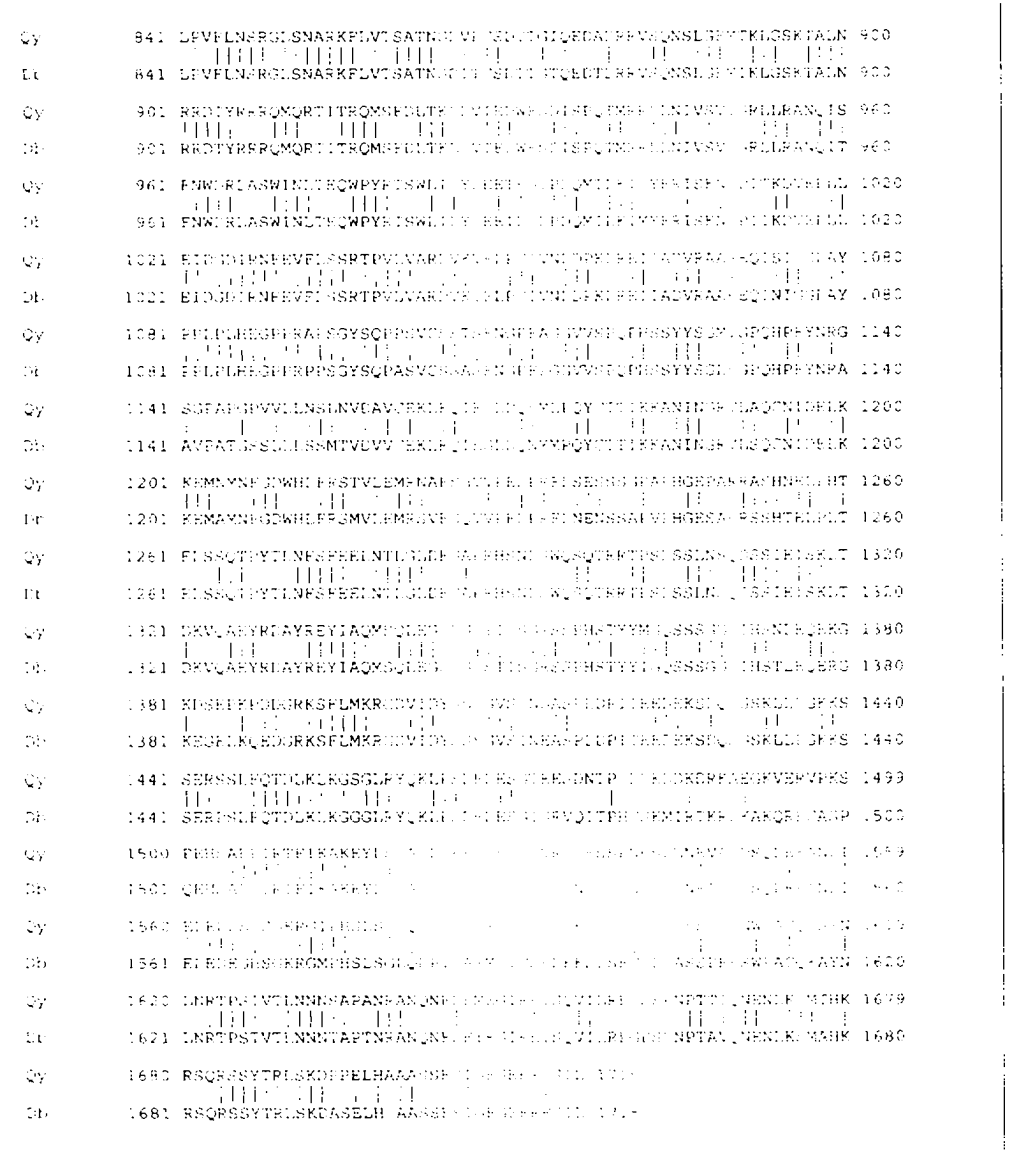
Db        661 FVIFLFIVGCIIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKHLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

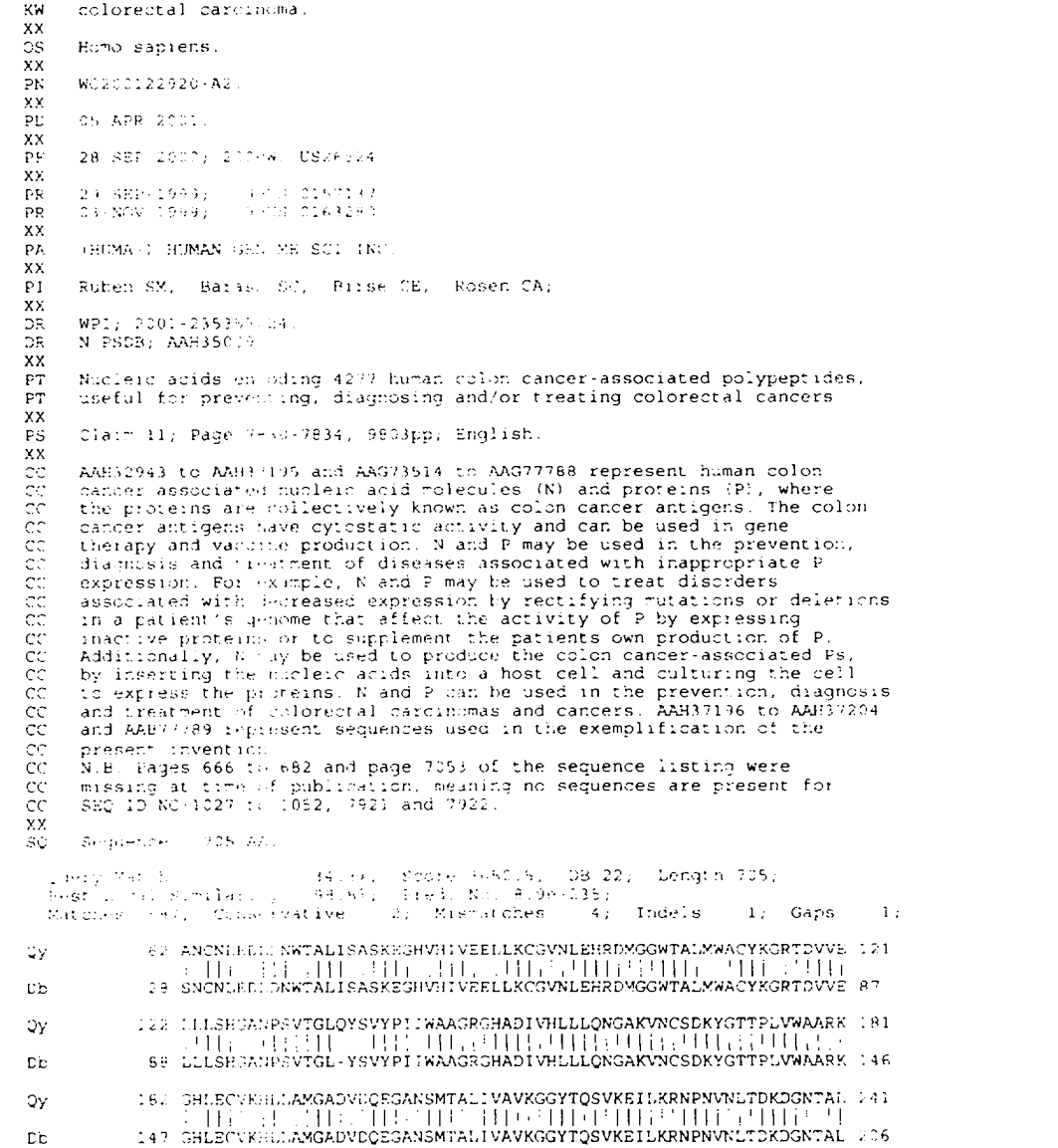
Db        721 NSQRKHLHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy        781 EQDKVLQVLDTVRVLFSKGFFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Db        781 EQDKVLQVLDTVRVLFSKGFFIAIFASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
```



RESULT 9
AAG75604
ID AAG75604 standard; Protein; 705 AA.
XX
AC AAG75604;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:1028.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis, [illegible]



KW colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05 APR 2001.
XX
PF 28 SEP 2000; 2000WO-US26224
XX
PR 24 SEP 1999; 1999US-0157147
PR 03 NOV 1999; 1999US-0163283
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235365/24.
DR N-PSDB; AAB35039.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers
XX
PS Claim 11; Page 7458-7834, 9903pp; English.
XX
CC AAB32943 to AAB37195 and AAG73514 to AAG77788 represent human colon
CC cancer associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with increased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAB37196 to AAB37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 705 AA;

  Query Match 3.[illegible]; Score 1[illegible]; DB 22; Length 705;
  Best Local Similarity 88.4%; Pred. No. 8.9e-035;
  Matches 1[illegible]; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```
Db        427 VSAVDKKGDTVKLHIAIRGRSRKLAELLRNHEKLGRDLYRPNKAGETPYNIDCSHQKSILT 486

Qy        422 QIFGARHLSPTETDGDMLGYDLYSEALADILSEPTMQPPITVGLYAQWI RKFLLKKLE 481
              |  ||||  ||||||  ||||   ||  ||  ||  ||  ||   ||  ||||
Db        487 QIFGARHLSPTETDGDMLGYDLYSEALADILSEPTMQPPITVGLYAQWJ RKFLLKKLE 546

Qy        482 DEMKTFAJQQIEFLFQFSWLIVFLLTLLLCGGLGLLFAFTVFGLIYVGLRHFITEVQIIQDS 541
              ||| ||||  ||||  |||  ||  ||  ||  ||  ||  ||  ||  ||||  ||
Db        547 DEMKTFAIQQIEFLFQFSWLIVFLLTLLLCGGLGLLFAFTVFGLIYVGLRHFITEVQIIQDS 606

Qy        542 IVIYFGGRRFGESWNWAWVLSTSLAHFIGYSIILGLMFVNFPELIPLI FALPVPFIFT 601
              ||  ||||  ||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||||  ||
Db        607 IVIYFGGRRFGESWNWAWVLSTSLAHFIGYSIILGLMFVNFPELIPLI FALPVPFIFT 666

Qy        602 DYNRLSSVGGETSLAEMIAILSDA PRFPGFIATRLRPVFKTFDTQGKF AKKT 654
              ||  ||||  ||||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db        667 DYNRLSSVGGETSLAEMIAILSDA PRFPGFIATRLRPVFKTFDTQGKF PKNS 719


RESULT 10
AAE01035
ID   AAE01035 standard; Protein; 551 AA
XX
AC   AAE01035;
XX
DT   04-JUL-2001 (first entry)
XX
DE   Human death domain containing receptor protein SEQ ID NO:56 protein from clone
XX
KW   Human; death domain-containing receptor; DDCR; immunosuppressive;
KW   antiarthritic; antirheumatic; antiproliferative; cytostatic;
KW   cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW   antibacterial; virucide; fungicide; ophthalmological; gene therapy;
KW   immunodeficiency disease; Acquired immune deficiency syndrome;
KW   AIDS; leukaemia; autoimmune disease; systemic lupus erythematosus;
KW   hyperproliferative disorder; neoplasm; cerebrovascular disorder;
KW   cerebral ischaemia; angiogenesis; cardiovascular disorder;
KW   neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW   ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;
KW   spinal muscular atrophy; epithelial cell proliferation; infection;
KW   cancer; wound healing; skin aging; chemotaxis; HOGDX59 clone.
XX
OS   Homo sapiens
XX
FH   Key             Location/Qualifiers
FT   Domain          49..616
FT                   /label= Transmembrane region
FT   Domain          123..144
FT                   /label= Transmembrane region
FT   Region          44..69
FT                   /label= Immunogenic epitope
FT   Region          131..136
FT                   /label= Immunogenic epitope
FT   Region          164..169
FT                   /label= Immunogenic epitope
FT   Region          197..203
FT                   /label= Immunogenic epitope
FT   Region          228..235
FT                   /label= Immunogenic epitope
FT   Region          298..316
FT                   /label= Immunogenic epitope
FT   Region          328..335
FT                   /label= Immunogenic epitope
FT   Region          445..452
FT                   /label= Immunogenic epitope
FT   Region          482..490
FT                   /label= Immunogenic epitope
FT   Region          540..551
FT                   /label= Immunogenic epitope
XX
PN   WO200129063-A2.
XX
PD   26-APR-2001.
XX
PF   17-OCT-2000; 2000WO-US28666.
XX
PR   18-OCT-1999;   99US-0159565.
PR   24-NOV-1999;   99US-0167246.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J;  Ruben SM;
XX
DR   WPI; 2001-300316/31.
DR   N-PSDB; AAD04779.
XX
PT   New death domain containing receptor polynucleotides and polypeptides,
PT   useful for treating and diagnosing cancer -
XX
PS   Claim 1; Page 29, 294; 298pp; English.
XX
CC   The patent discloses novel death domain-containing receptor (DDCR)
CC   cDNAs and their corresponding proteins. DDCR cDNA and protein are
CC   used to prevent, treat or ameliorate a medical condition in mammals.
CC   They are also used in diagnosing a pathological condition or
CC   susceptibility to a pathological condition. The DDCR protein and
CC   its antibodies are used in the diagnosis and treatment of disorders
CC   such as immunodeficiency diseases (e.g. Acquired immune deficiency
CC   syndrome (AIDS), leukaemia) autoimmune diseases (e.g. systemic lupus
CC   erythematosus, rheumatoid arthritis), hyperproliferative disorders
CC   (e.g. neoplasms of the breast or liver), cerebrovascular disorders
CC   (e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders
CC   (e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's
CC   disease, Parkinson's disease), ocular disorders (e.g. corneal
CC   infection), degenerative diseases (e.g. spinal muscular atrophy-SMA),
CC   cancer, aberrant apoptosis, disorders of the placenta or uterus and
CC   infections caused by bacteria, viruses and fungi. The DDCR proteins
CC   are used to aid wound healing and epithelial cell proliferation,
CC   to prevent skin aging due to sunburn, to maintain organs before
CC   transplantation, to support cell culture of primary tissues, to
CC   regenerate tissues and in chemotaxis. The DDCR sequences are used
CC   in correcting aberrant cellular apoptosis by gene therapy.
CC   The present sequence is human death domain-containing receptor
CC   (DDCR) protein from clone HODDX59.
XX
SQ   Sequence   551 AA;

  Query Match          32.1%;  Score 2851.5;  DB 22;  Length 551;
  Best Local Similarity   99.6%;  Pred. No. 5.1e-319;
  Matches  550;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        102 MGGWTALVWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNG 161
              |  |||  |||| |||| |||  |||| ||||||||||||||||||||||||||||||||
Db          1 MGGWTALVWACYKGRTDVVELLLSHGANPSVTGL-YSVYPIIWAAGRGHADIVHLLLQNG 59

Qy        162 AKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVK 221
              | ||||  |||| |||| |||| |||| ||||||||||||||||||||||||||||||||
Db         60 AKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVK 119

Qy        222 EILKRNPNVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 281
              | ||||  |||| |||| |||| |||| |||| |||||||||||||||||||||||||||
Db        120 EILKRNPNVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 179

Qy        282 HVEIVPALQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 341
              |  |||  ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
Db        180 HVEIVPALQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLI 239

Qy        342 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRP 401
              |  |||  ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
Db        240 KATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRP 299

Qy        402 NKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 461
              |  |||  ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
Db        300 NKAGETPYNIDCSHQKSILTQIFGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 359

Qy        462 GVGLYAQWIRKFLLKKLEDEMKTFAGQQIEPLFQFSWLIVFLTLLLCGGLGLLFAFTV 521
```

```
Db        360 CVGLVACWCDGKSELLKKLDECEMKTPA[illegible]PQPSWIYSETLDETGDGIEAFTV 419

Qy        522 RENLGIAVRLSFLALLYIFIVIYFIAEFI[illegible]WRWAWVLSTFIAAGI[illegible]LEDLIFIVEV 601

Db        420 RPNLGIAVRLSFLALLYIFIVIYFIAEFI[illegible]WRWAWVLSTF-AAGI[illegible]LEDLIFIVEV 498

Qy        602 NPPELIEQTIRALIVPFFIFIYNFLYCIEIL[illegible]AAPMIATIIVFA[illegible]LATPIFAVF 681

Db        499 NPPELIEQTIRALIVPFFIFIYNFLYCIEIL[illegible]AAPMIATIIVFA[illegible]LATPIFAVF 578

Qy        682 KTETDDKPKWK 693
              ||||| |||
Db        579 KTETDDKPKWKKK 591


RESULT 11
AAB95191
ID   AAB95191 standard; Protein; 543 AA.
XX
AC   AAB95191;
XX
DT   26-JUN 2001 (first entry)
XX
DE   Human protein sequence SEQ ID NO:[illegible]
XX
KW   Human; primer; detection, diagnosis, [illegible] therapy.
XX
OS   Homo sapiens
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; [illegible]
XX
PR   29-JUL-1999;   99EP-0401683
PR   27-AUG-1999;   99JP-0241225
PR   11-JAN-2000;   2000JP-0001877
PR   08-MAY-2000;   2000EP-0303763
PR   09-JUN-2000;   2000JP-0174894
XX
PA   (HELI ) HELIX RES INST.
XX
PI   Kita T, Ishii S, Nishikawa [illegible]
PI   Ishii S, Okuyama T, Wada [illegible]
XX
DR   WPI; 2001-[illegible]
XX
PT   Primer set for synthesis [illegible]
PT   full length cDNA obtained [illegible]
PT   and/or diagnosis of the [illegible]
PT   full length cDNAs.
XX
PS   Claim 8; SEQ ID 543; 254pp; [illegible]
XX
CC   The present invention describes [illegible] full-length cDNAs defined in the [illegible]
CC   full-length cDNAs defined in the [illegible] which a primer set
CC   comprises: (a) an oligo-dT primer [illegible] complementary
CC   to the complementary strand of a [illegible] and (b) [illegible] one of
CC   the 5602 nucleotide sequences described in the specification, where the
CC   oligonucleotide comprises at least [illegible] combination
CC   of an oligonucleotide comprising a [illegible] the
CC   complementary strand of a polynucleotide and [illegible] comprises a first
CC   sequence and an oligonucleotide [illegible] complementary to a
CC   polynucleotide which comprises a [illegible] sequence, wherein the
CC   oligonucleotide comprises at least [illegible] of the nucleotide at
CC   the 5' end sequence/3'-end sequence [illegible] described in
CC   the specification. The primer sets [illegible] therapy and
CC   in gene therapy. The primers are [illegible] synthesizing [illegible],
CC   particularly full length cDNAs. [illegible] are also useful for the
CC   detection and/or diagnosis of the [illegible] and
CC   the full length cDNAs. The [illegible]
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH17633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   543 AA;

  Query Match           81.4%;  Score 2791;  DB 22;  Length 543;
  Best Local Similarity   99.6%;  Pred. No. 3.5e-214;
  Matches  541;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1171 MLPQYMTTIKKANTNGRVLAQCNITELKKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDP 1232
              ||||  ||    |||  |||  ||| | |||| |||| |||| ||||| |||| | |||
Db          1 MLPQYMTTIKKANTNGRVLAQCNITELKKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDP 60

Qy       1233 RFLSESSGIAPHGEPARRASHNELPHTELSSQTPYTLNFSPEELNTLGLDEGAPRHSNL 1292
              |||||.||    |||  |||  ||| | ||| |.|||| ||||| |||| |||| | | |
Db         61 RFLSESSGIAPHGEPARRASHNELPHTELSSQTPYTLNFSPEELNTLGLDEGAPRHSNL 120

Qy       1293 SWQSQTRSTISLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISG 1352
              |||| |||| |||  ||| |||||| |||| |||||| |||||||||| |||| | |||
Db        121 SWQSQTRSTISLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISG 180

Qy       1353 RSSPRHTYYMGQSSSGGSIHSNLEQEKGKDSEPKPDDGRKSFLMKRGDVIDYSSSGVSTN 1412
              ||| |||| ||| |||||| |||| |||||||||| |||| |||||| |||| |||
Db        181 RSSPRHTYYMGQSSSGGSIHSNLEQEKGKDSEPKPDDGRKSFLMKRGDVIDYSSSGVSTN 240

Qy       1413 DASPLGMFTEEDEKSDQSGSKLLDGKKSSERSSLFQTDLKLKGSGLRYQKLPSDEDFSGT 1472
              | |||  |||  |||  ||||  |||| |||| |||||  |||||  |||| |||| |||
Db        241 DASPLGMFTEEDEKSDQSGSKLLDGKKSSERSSLFQTDLKLKGSGLRYQKLPSDEDESGT 300

Qy       1473 EESDNTPSLKDDKDRKAESGKVERVFKSPEHSAEPIRTFIKAKEYLSDALLDKKDSSDSGV 1532
              ||||  ||    |||  |||  .||  |||| |||| |||| |||| |||| |||| |||
Db        301 EESDNTPSLKDDKDRKAESGKVERVFKSPEHSAEPIRTFIKAKEYLSDALLDKKDSSDSGV 360

Qy       1533 RSSFSSPHHSLHNEVADDSQLEKANLIELEDDSHSGKRGIPHSLSGLQDPIIARMSICSE 1592
              ||| | || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db        361 RSSESSPHHSLHNEVADDSHLEKADLIELEDDSHSGKRGIPHSLSGLQDPIIARMSICSE 420

Qy       1593 DKKSEIPSLIASSPEENNPACQKAYNLNRTPSTVTLNNNSAPANRANQNFDEMEGIRET 1652
              |||  ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db        421 DKKSEIPSLIASSPEENNPACQKAYNLNRTPSTVTLNNNSAPANRANQNFDEMEGIRET 480

Qy       1653 AQVIPSIPAPNPTTIQNENLKSMIHKPSQRSSYTRLSKDPPELHAAASSESTGFGEERE 1712
              |  |||  ||| |||  ||| |||| |||| |||| |||| |||| |||| |||| |||
Db        481 AQVIPSIPAPNPTTIQNENLKSMIHKPSQRSSYTRLSKDPPELHAAASSESTGFGEERE 540

Qy       1713 SIS 1715
              |||
Db        541 SIS 543


RESULT 12
AAB03648
ID   AAB03648 standard; Protein; 512 AA.
XX
AC   AAB03648;
XX
DT   06-AUG-2001 (first entry)
XX
DE   Human extracellular matrix and cell adhesion molecule-9 (XMAD 9).
XX
KW   Human; extracellular matrix and cell adhesion molecule; XMAD;
KW   gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
KW   Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
KW   sickle cell anaemia; thalassaemia, autoimmune disorder; adenocarcinoma;
KW   inflammatory disorder; acquired immune deficiency syndrome; AIDS;
KW   Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma,
KW   Crohn's disease; diabetes mellitus, atopic dermatitis; lymphoma; cancer;
KW   glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
KW   osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis,
KW   infection; cell proliferative disorder; actinic keratosis, myeloma;
```

```
KW   arteriosclerosis; nootropic; anticonvulsant; antirheumatoid; nephrotropic;
KW   neuroprotective; dermatological.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          94..108
FT                   /note= "Adhesion-associated domain"
FT   Domain          196..199
FT                   /note= "RGD cell attachment motif"
FT   Region          328..445
FT                   /note= "Plasma fibronectin family"
XX
PN   WO200142288-A.
XX
PD   14-JUN-2001.
XX
PF   05-DEC-2000; 2000WO-US32993.
XX
PR   10-DEC-1999;   99US-0172852
PR   16-DEC-1999;   99US-0172454
XX
PA   (INCY ) INCYTE GENOMICS INC.
XX
PI   Yue H, Tang YT, Lal P, Rolfe M, Azimzai Y, Gattiker C;
PI   Baughn MR, Emerling BM, Shah P, Azimzai Y;
XX
DR   WPI; 2001-281A1L/40.
DR   N-PSDB; AAG02064.
XX
PT   New human extracellular matrix and cell adhesion molecules and
PT   polynucleotide sequences encoding them, useful for diagnosis,
PT   prevention, treatment of genetic, inflammatory and cell proliferative
PT   disorders
XX
PS   Claim 1; Page 191-192; 196pp; English.
XX
CC   The present sequence is a human extracellular matrix and cell
CC   adhesion molecule (XMAD). The XMAD is used for screening a compound for
CC   effectiveness as an agonist or antagonist. XMAD, the specified agonist
CC   or antagonist are used for treating a disease or condition associated
CC   with decreased or increased expression or function of XMAD. The
CC   polynucleotides encoding XMAD are useful in somatic or germ line gene
CC   therapy to correct a genetic deficiency, to express a gene normally
CC   lethal gene product and to express [illegible] sufficient protection
CC   against intracellular [illegible]. [illegible]
CC   associated with expression of XMAD, [illegible] hybridization
CC   hybridisation probes [illegible]. [illegible]
CC   sequences and to create [illegible]. [illegible]
CC   animals [illegible]. [illegible]
CC   fragments derived from the polynucleotides may be used as
CC   elements on a microarray. Antibodies which specifically bind XMAD may be
CC   used for the diagnosis of disorders associated with the expression of
CC   XMAD, or in assays to monitor patients being treated with XMAD. Diseases
CC   diagnosed, prevented or treated include genetic disorders such as
CC   adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
CC   disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
CC   autoimmune/inflammatory disorders [illegible] immune deficiency
CC   syndrome (AIDS), Addison's disease, alopecia, anaemia, anorexia,
CC   atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
CC   glomerulonephritis, multiple sclerosis, Graves' disease, osteoarthritis,
CC   osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC   bacterial, fungal, parasitic, protozoal and helminthic infections and
CC   cell proliferative disorders such as [illegible] metastasis, atherosclerosis
CC   and cancer including breast, bladder, [illegible] brain, [illegible] uterus
CC   cancer, leukaemia, adenocarcinoma, [illegible] and [illegible]oma
XX
SQ   Sequence   513 AA;

  Query Match             28.7%;  Score [illegible];  DB 21;  Length 513;
  Best Local Similarity   91.8%;  Pred. No. [illegible];
  Matches  512;  Conservative    0;  Mismatches   [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy       1203 MNMNPGDWHLPRSTVLEMSNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 1262
             ||| | |  |||| ||  | |||| |||| |||| |||||||||| ||||
Db          1 MNMNPGDWHLPRSTVLEMSNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 60

Qy       1263 SSQTPYTLNPSPREDLNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDK 1322
              |||| | | |||  |||| |||| ||||||||||| |||||||||||| |||| ||||
Db         61 SSQTPYTINPSPREDLNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDK 120

Qy       1323 VQAEYRTAYRFYIAQMSQLEGSPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKGKD 1382
              |||  |||  |||  |||  |||| |||||||||||| ||||  |||| |
Db        121 VQAEYRTAYRFYIAQMSQLEGSPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKGKD 180

Qy       1383 SEPKPDP PKSFLMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE 1442
                 ||   || ||| |||  ||| |||  |||| | |||| | |||| ||||  ||||
Db        181 SEPKPDPAPKSFLMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE 240

Qy       1443 RSSLPQTPLKLKGSGLRYQKLPSGEDESGTEESDNTPLLKDDKDRKAEGKVERVPKSPEH 1502
              ||||  | | |||| |||  |||  |||| |||| |||| |||| ||||
Db        241 RSSLPQTPIKLKGSGLRYQKLPSGEDESGTEESDNTPLLKDDKDRKAEGKVERVPKSPEH 300

Qy       1503 SAEPIPDPIKAKEYLSDALLDKKDSSDSGVRSSESSPNHSLHNEVADDSQLEKANLIELE 1562
              | |||  |||| ||| |||| |||| |||| |||| |||| |||| ||||
Db        301 SAEPIPDPIKAKEYLSDALLDKKDSSDSGVRSSESSPNHSLHNEVADDSHLEKANLIELE 360

Qy       1563 DLSHSGRPGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNR 1622
              | |||  ||| |||| |||| |||| |||| |||| |||| |||| ||||
Db        361 DDSHSGRPGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNR 420

Qy       1623 TPSTVTLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRSQ 1682
              || |||  ||| ||| |||  |||| |||| |||| |||| |||| ||||
Db        421 TPSTVTLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRSQ 480

Qy       1683 RSSYTPLNFLPPELHAAASSESTGPGEERESIL 1715
              || |||  |||| ||| |||| |||| ||
Db        481 RSSYTPLNFLPPELHAAASSESTGPGEERESIL 513


RESULT 13
ABG08697
ID   ABG08697 standard; Protein; 624 AA.
XX
AC   ABG08697;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Novel human diagnostic protein #08794.
XX
KW   human; diagnostic, mapping, gene mapping, gene therapy; forensic;
KW   full length, cytoplasmic, signal peptide; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200176067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   22-AUG-2000; 2000US-0643167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639842/73.
DR   N-PSDB; AAS72884.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity
XX
```

```
PS   Claim 20; SEQ ID No 39056; 107pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and in chromosome
CC   and gene mapping, and in recombinant production of (II). (I) and
CC   polynucleotides are also used in diagnostic assays as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II). (I) is useful for treating disease states involving
CC   (II). (II) is useful for generating antibodies against it, determining or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (III) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful in treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010 ABG01177 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence   644 AA;

  Query Match             28.2%;  Score 1183;  DB 21;  Length 644;
  Best Local Similarity   32.2%;  Pred. No. 1.2e-111;
  Matches  497;  Conservative  301;  Mismatches  544;  Indels   61;  Gaps    5;

Qy       1183 KANINGRVLAQCNIDELKKEMDNLA-WDLFS-TVLEMPNSAHHVVEL--RPLSESSSGP 1242
             :||   .||:  :|||   |:|          :  | :    :    ||   ||  :
Db         86 EANINGRVLAQCNIDELKKEMDNLA-WDLFS-TVLEMPNSAHHVVEL--RPLSESSSGP 145

Qy       1243 APHGEIARRASHNELPHTELSQCTYETDPVFG-NTETHEVAPHHNL-SWQSQIPRTP 1300
             ||||| : |||  |:|  :||  :  :  :  |  |||  |||  |  |||  ||
Db         146 APHEEIARRASHNELPHTELSQCTYETDPVFGKNTETHEVAPHHNLESWQSQIPRTP 205

Qy       1301 SLSQNSQDSSIEISKLTDKVQAEFSTGLYQFESEIDECDHSPHPHSYYM 1362
             |:|  :|||:  :|||:  :|||  :  :   :|  :|  :|  :  |||  ||
Db         206 SLSYNSQDSSIEISKLTFKVQAEFSTGLYQFESEIDECDHSPHPHSYYM 265

Qy       1363 GQASSQNSIHSNLAQEPQYDNEPVL-RRPSVMAPSQVIYSKLGKPIT-ASPLI-ITE 1420
             |:|  :|||:  :|||  :|:  :  :   :|  :|  :|  :  |||  ||
Db         266 GQASSQNSIHSNLAQEPQYDNEPVL-RRPSVMAPSQVIYSKLGKPIT-ASPLI-ITE 323

Qy       1421 E[illegible]                                    K 1462
             :|                                               :
Db         324 E[illegible]                                    K 365

Qy       1463 DLKSYASQKVESVPPSSIHSAFS-----PQFSNSSLSVYSTKSEASFN 1510
             :||  |  |||  :||  :  :  :     :  :  :  :  :  :  :
Db         366 DDKSRKPEQKVERVPPSSIHSAVS-----PAFSNSSLSVYSTKSEASFN 415

Qy       1511 H-SLHNEVADD SQLEKANLIPLPIP-IPSLPSLPSLPSLPSLPSLPSLPSLPSLPS 1568
              |  :  :||   |::   :|  :  :  :  :  :  :  :  :  :  :  :
Db         416 HVSAQCKQADDCPDLKRHNLIRLPIP-IPSLPSLPSLPSLPSLPSLPSLPSLPS 475

Qy       1569 EGSLI-ASSPEENWPAQQKAYNQPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIP 1626
              :  :  :|||  :|  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         476 RIASLQKPSSPEENWPAQQKAYNQPIPIPIPIPIPIPIPIPIPIPIPIPIPIPIP 535

Qy       1627 LRPSGSPNPTTIQNENQKSMTIKP-QPPSSPIPIPIPIPIPIPIPIPIPIPIPIP 1685
              :||  :|||  :|||  :||  :  :  :  :  :  :  :  :  :  :  :
Db         536 LRPSGSPNPTTIQNENQKSMTIKP-QPPSSPIPIPIPIPIPIPIPIPIPIPIPIP 594


RESULT 14
AB564857
ID   AB564857 standard; Protein; 1498 AA.
XX
AC   AB564857;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 21363.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171041 A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US08281.
XX
PR   23-MAR-2000; 2000US-0191679P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC;  Adams M;  Li PWD;  Myers EW;
XX
DR   WPI; 2001-656845/75.
DR   N-PSDB; AB038900.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 21363; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL06176-ABL30511), expressed DNA
CC   sequences (ABL01440-ABL16175) and the encoded proteins
CC   (AB857727-ABB00000).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1498 AA;

  Query Match             26.8%;  Score 1124.5;  DB 22;  Length 1498;
  Best Local Similarity   32.7%;  Pred. No. 2e-108;
  Matches  544;  Conservative  299;  Mismatches  490;  Indels  179;  Gaps   39;

Qy          1 YSVLL-L-VINYVSEENIPALKALLE-KCKDVEERNECGQTPLMIAAEQGNLEIVKELIK 59
              |  :  :  :  |  :  :  :  :|:|:|  :  |  :| :  :    :  :
Db         23 MGSLAIRAALQYIDNNISGLRAILDSRHLTIDDRDENATTVLMVVAGRGLTAFVREFLA 82

Qy         60 NGANCNLEDLDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDV 119
              :   |:| :|||:  |:  :  :::  ||    :|||| ||||:|||  |:|  |:
Db         83 RGADVLAEDLDNWTALLCASRNGHLDVVQLLLDHGAEVEHRDMGGWTSLMWAAYRGHTEL 142

Qy        120 VELLL-E-ANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAA 179
              |  :  :  :  |  :  :|:|  ||:  :||  :  :|  |  :|  :  |||  :
Db        143 VRLLLLR-ADGNAHG-NYHLGALLWAAGRGYKDIVELLVQRGAKVNVGDKYGTTALVWAC 201

Qy        180 RKGHLEIVKHLLAMGADVDCEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNT 239
              |:  :  :  |:  ||:  :  |  |  |:||  :  :|  |||  |||| :
Db        202 RRGNVEIVDTLLKAGANVDTAGMYSWTPLLVAAAGGHTDCVSSILEKKPNVNALDKDGMT 261

Qy        240 ALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIR 299
              :|  :|  :|  |:  :|  :|  :|  :|  :  :  :  :  :  :  :  :
Db        262 ALCIASREHFQDIAASLIAAGAYINIQDRGADTPLIHAVKAGHRTVVEALLKKHADVDIQ 321

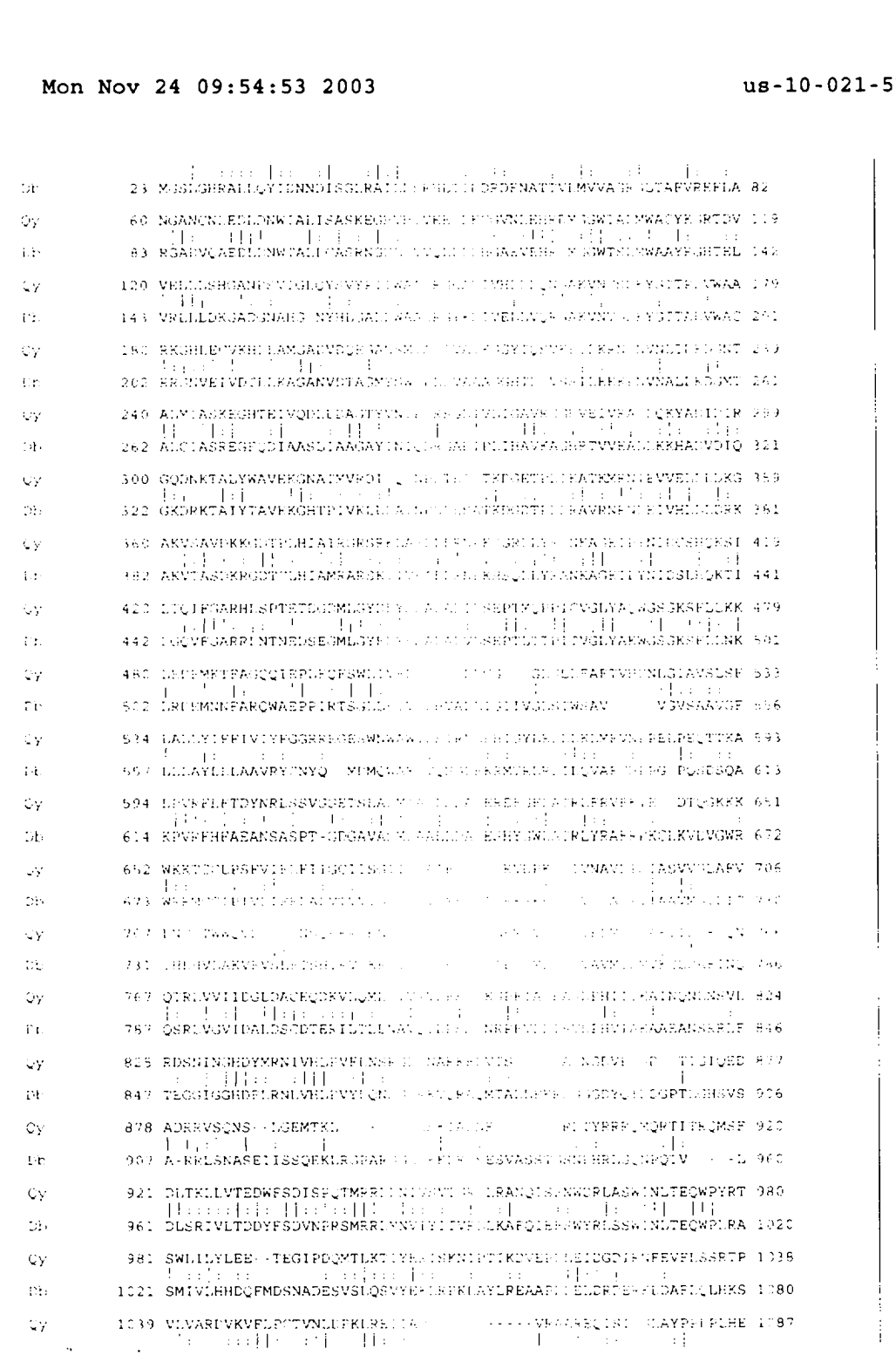
Qy        300 GQDNFIALTWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKG 359
              :|  :|  :|||:  :|:  :|  |  |||:|||:  :  :  :  :  :  :
Db        322 GKDRFALYWAVEKGHIPIVKLLIAINPDLESATKDGDTPLLRAVRNRNLEIVHLLLDRK 381
```

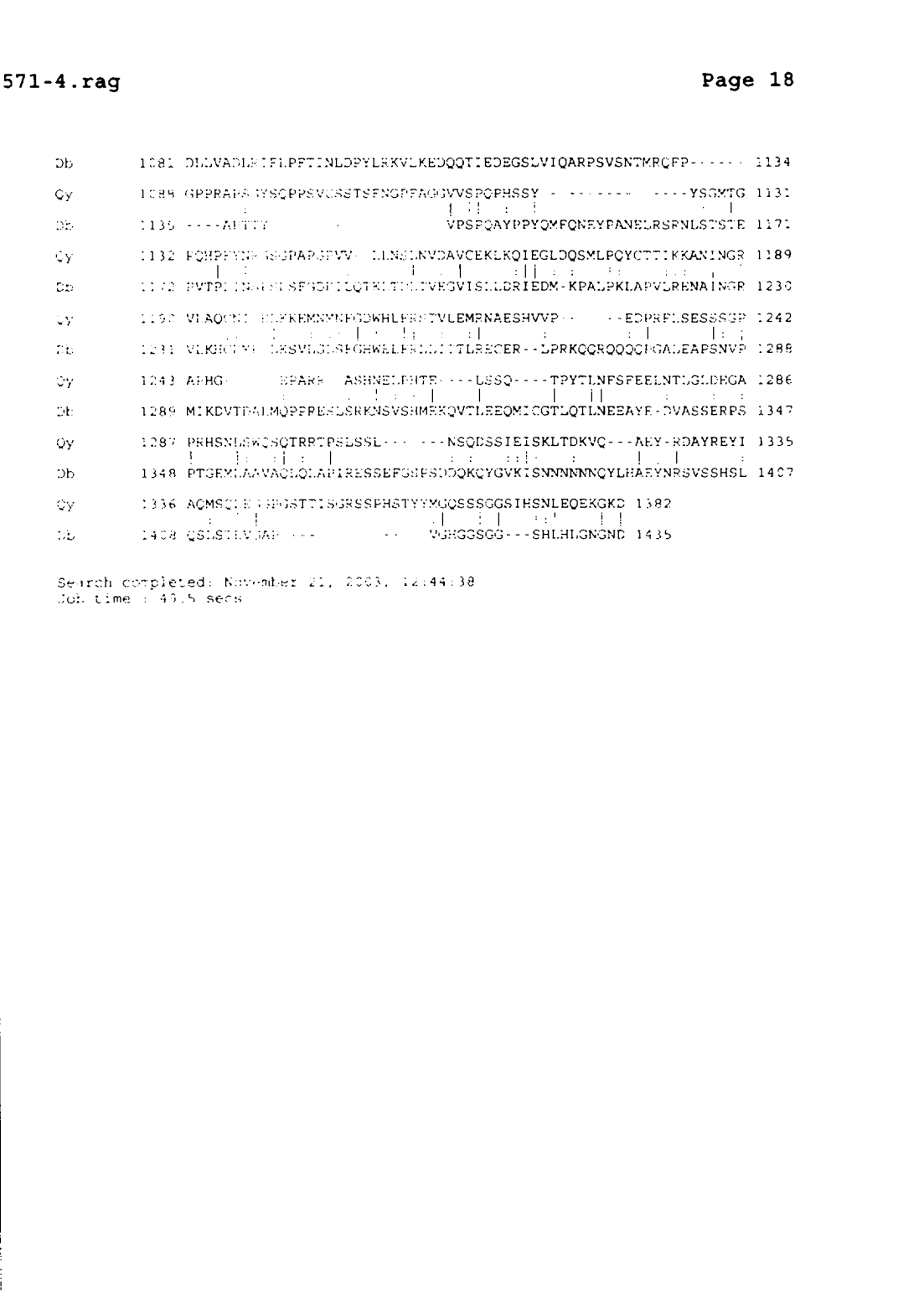
```
Qy    360 AKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSI 419
Db    382 AKVTASDKRGDTCLHIAMRARSKTIVEALLRNKHSQLLYRANKAGETPYNIDSLHQKTI 441
Qy    420 LTQIFGARHLSPTETGDMLGYDLY ALAD SEPTMQPPIGVGLYAQNGSGKSFLLKK 479
Db    442 LGQVFGARRLNTNEDSEGMLGYELY ALAD SEPTTTPITVGLYAKNGSGKSFLLNK 501
Qy    480 LEDEMKTFAGQQIEPLFQFSWLIV   FAFTVH NLSIAVSLSF 533
Db    502 LFDEYNNFARQWAEPPIRTSGLFIV HVALLGTIVHLATWIAV  VSVSAAVGF 556
Qy    534 LALLYIFFIVIYFNGRREGESWNWAWL STPLARHIGYLEI KLMPVN PELPEQTIKA 593
Db    557 LLLAYLLLAAVRY NYQ MDMQWAY VQHDL KRMTKIRLIIQVAPGH PG PQSDSQA 613
Qy    594 LPVRFLFTDYNRISSVGGETSLAEYIAIISDAVRREFGFIAT LERVPKIE DTQQKKK 651
Db    614 KPVRFHFAEANSASPT-GGGAVAHMLAALLDAIESHYGWIATRLYRAFR KCLKVDVGWR 672
Qy    652 WKKTCCLPSPVIFLFIIGCIISGIILAIF    RVDPKHLTVNAVLIIASVVGLAFV 706
Db    673 WRRMCCIPIVLIFELALVTVVTGISIVAYFTLAFEKEKEHIIV ALYLAAVMITLIC 730
Qy    707 LNCRTWWQVIDSLLNSQRKRLENAASF HKI LGEMKVLK VELMARYAKTIPSPTQN 766
Db    731 THLHVLAKVFVSLFTSHIRVLKRAY    FGS APLTMLGAEVAVMTIYVKCLSAFTNQ 786
Qy    767 QTRLVVIIDGLDACEQDKVLQMLDTV VLFS  KTPPIAIFASDPHIIIKAINQNLNSVL 824
Db    787 QSRLVGVIDALDSCDTERILTLLNAV LLL  NRPPVLIISVDPHVIAKAAEANSRRLF 846
Qy    825 RDSNINGHDYMRNIVHLPVFLNSR  NARFF VIS   ATLIDVPE  TTGIQED 877
Db    847 TEGIIGGHDFLRNIVHLPVYLQNS LAFVQSALYTALLFKRSDGDYQTLDGHTLGHSVS 906
Qy    878 ADRRVSQNS--LGEMTKL       SKTALNR      EDIYRRRQVQRTITQMSF 920
Db    907 A RFLSNASEIISSQFKLRGPARGL KFLRLSESVASST NN HRLSQSHQTV-----L 960
Qy    921 DLTKLLVTEDWFSDISPQTMRRLLNIVSVTGR LFANQISFNWKRLASW NLTEQWPYRT 980
Db    961 DLSRIVLTDDYFSDVNPRSMRRLMNVIYITVRILKAFQIEFSWYRLSSW NLTEQWPLRA 1020
Qy    981 SWLIDYLEE--TEGIPDQMTLKTIY      TTKRVEPI LDGDIPNEDVFLSSRTP 1038
Db   1021 SMIV        AAFDSL        VNL  NIS  TAVL  RS 1080
Qy   1039 VLVAKLVEVLLDAVNIQ FLSP                     AYI    1087
Db   1081 DLLVADLRIFLPFTINLDFTGKKVL                    VSANI  QIP 1134
Qy   1088 GPPKAPSGYSQPPSVCSSTSFNGPPA VVAL       --YSGMTG 1131
Db   1135 ----APTTY                   VIS LAYPPYQMEQN YPANL VSRNLSTSTE 1171
Qy   1132 PQHPFYNRGSGPAPGPVV  LLNSLNLKAV    KQISDLHQS LPQYGIIKKANINGR 1189
Db   1172 PVTPLINSPSDSFGDDILQTKLTDL LKE   IGLIPRIEEL KLALPKLAIVLRENAINGR 1230
Qy   1190 VLAQCNIDELKKEMNMNFGDWHLFL IVLEMSNARSPVTF        HEDSPFLSESSSGP 1242
Db   1231 VLKHCDMPDLKSVLGLSFGHWELFRILTTIAKTER  LPPKQLRQQQQHALEAFSNVP 1288
Qy   1243 APHG   --EPARR--ASHNELPHTE   LNSQ  TPYIGFSPEEINTLGLDEGA 1286
Db   1289 MIKDVTDALMQPPRESLSRKNSVSHMEALVILEEQMI GTLQTENEEAYS DVASSERPS 1347
Qy   1287 PRHSNLSWQSQTRRTPSLSSL--      NSQPSSIEISKLTDPKVQ  AL  RDAYREYI 1335
Db   1348 PTGEMLAAVAQLQLAPIRESSEFGSPSQQKQYGVKISNNNNNNQYLHA NPAVPHSL 1407
Qy   1336 AQMSQLEGGPGSTITISGRSSPHSTYYM LSS  SIHSNLPQEPGKD  42
Db   1408 QSLSTLVGAP--------        VGHGGSGG---SHLHLGNGND 1435
```

RESULT 15
AAE25146
ID   AAE25146 standard; Protein; 1498 AA.
XX
AC   AAE25146;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Fruit fly ARMS protein.
XX
KW   Ankyrin repeat rich membrane spanning protein; ARMS; neurotrophin;
KW   ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW   diagnostic; imaging; fruit fly.
XX
OS   Drosophila sp.
XX
FH   Key   Location/Qualifiers
FT   Region   476..515
FT   /note= "N-terminal region"
FT   Region   594..657
FT   /note= "Loop between TM2 and TM3"
FT   Region   740..848
FT   /note= "Carboxy tail"
FT   Region   480..1102
FT   /note= "Carboxy tail"
FT   Domain   1193..1255
FT   /note= "SAM domain"
FT   Binding-site   1496..1499
FT   /note= "PDZ binding motif"
XX
PN   WO200250273-A2.
XX
PD   27-JUN-2002.
XX
PF   19-DEC-2001; 2001WO-US48603.
XX
PR   21-DEC-2000; 2000US-256909P.
XX
PA   (UYNY) UNIV NEW YORK STATE.
XX
PI   Chao MV, Kong H;
XX
DR   WPI; 2002-508854/54.
DR   N-PSDB; AAD41034.
XX
PT   Isolated ankyrin repeat rich membrane spanning (ARMS) polypeptide that
PT   is a target for phosphorylation by neurotrophin and ephrin receptor
PT   tyrosine kinases, useful as a marker for growth cones -
XX
PS   Disclosure; Page 130-136; 136pp; English.
XX
CC   The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
CC   protein which is a down stream target of neurotrophin and ephrin
CC   receptor tyrosine kinases. ARMS DNA is useful for visualising the
CC   growth cone of neurons. ARMS protein is useful as an indicator of
CC   the biological activity of neurotrophins and ephrins, as a marker
CC   for neuronal cells which have the ability to undergo continued
CC   synaptic changes through adult life or for the presence and
CC   distribution of ARMS in such neuronal cells, or as a marker for
CC   growth cones. ARMS DNA is useful in diagnostic and imaging methods.
CC   The present sequence is fruit fly ARMS protein.
XX
SQ   Sequence   1498 AA;

Query Match 26.8%; Score 2379.5; DB 23; Length 1498;
Best Local Similarity 37.6%; Pred. No. 2e-180;
Matches 559; Conservative 259; Mismatches 490; Indels 179; Gaps 39;

```
Qy      1 MSVLISQSVINYVEEENIPALKALLE KCKDVDERNECGQTPLMIAAEQGNLEIVKELIK 59
```





Search completed: November 21, 2003, 12:44:38
Job time : 40.5 secs

OmniSeq version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 12:34:[illegible] ; Search time 8.28 seconds
(without alignments)
9[illegible].8474 Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8884
Sequence: 1 MSVLISQSVINYVERENIPV[illegible] LRAAAA[illegible]TDFGSRsdIL 1777

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 8[illegible]525 seqs, 2567[illegible]604 residues

Total number of hits satisfying chosen parameters: 8[illegible]525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_[illegible]:*
17: sp_[illegible]:*

Pred. No. 1 [illegible]
score greater than no equal to [illegible]
and is derived by analysis of [illegible]

SUMMARIES

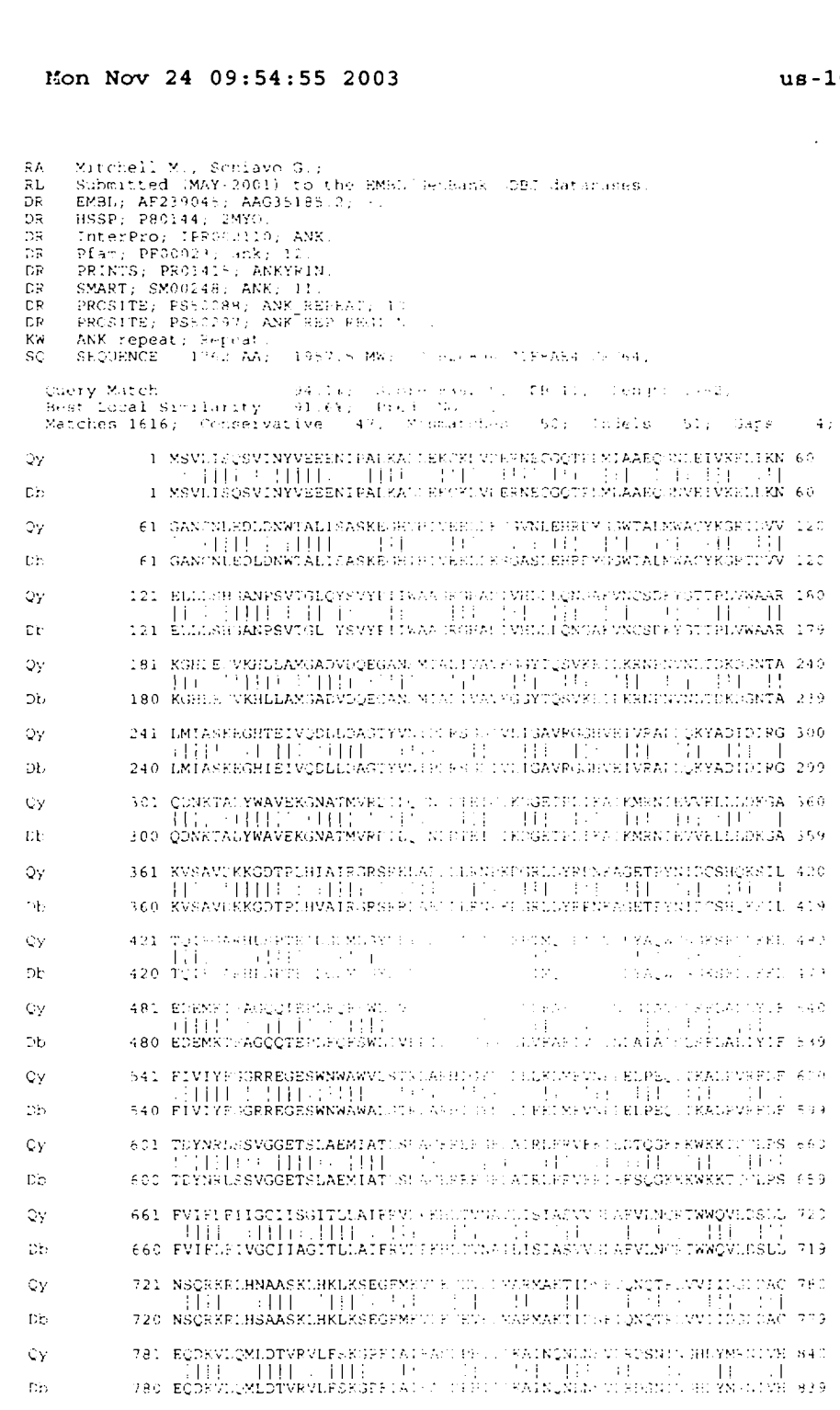
Result No.	Score	% Query Match	Length	DB	ID	Description
1	8860	99.4	1777	4	Q9UJH0	Q9UJH0 homo sapien
2	8562.5	96.1	1762	11	Q9EL[illegible]	Q9EL[illegible] rattus norv
3	8266	93.0	1715	11	Q9EP[illegible]	Q9EP[illegible] rattus norv
4	5205.5	58.6	1031	4	Q9U[illegible]	Q9U[illegible] homo sapien
5	2791	31.4	543	4	Q9[illegible]	Q9[illegible] homo sapien
6	2379.5	26.8	1479	5	Q8W[illegible]	Q8W[illegible] drosophila
7	1832.5	20.6	1398	5	Q22[illegible]	Q22[illegible] caenorhabdi
8	1255	14.1	246	4	Q9NT[illegible]	Q9NT[illegible] homo sapien
9	1119.5	12.6	239	11	Q8K[illegible]	Q8K[illegible] mus musculu
10	666	7.5	129	4	Q8WA[illegible]	Q8WA[illegible] homo sapien
11	657.5	7.4	1858	4	Q994[illegible]	Q994[illegible] homo sapien
12	644	7.2	426	16	Q8YI[illegible]	Q8YI[illegible] anabaena sp
13	623	7.0	1943	11	Q61[illegible]	Q61[illegible] mus musculu
14	622.5	7.0	1726	11	Q4V[illegible]	Q4V[illegible] mus musculu
15	622	7.0	1713	4	Q14[illegible]	Q14[illegible] homo sapien
16	617	6.9	2682	11	Q[illegible]	Q[illegible] rattus norv
17	606	6.8	1849	11	Q61302	Q61302 mus musculu
18	600.5	6.8	6994	5	Q17343	Q17343 caenorhabdi
19	600.5	6.8	6994	5	Q17490	Q17490 caenorhabdi
20	498.5	5.6	1136	6	Q9N180	Q9N180 bos taurus
21	496.5	5.6	1219	11	Q8C8R3	Q8C8R3 mus musculu
22	495.5	5.6	1159	5	Q9NCP8	Q9NCP8 drosophila
23	493.5	5.6	844	11	P97582	P97582 rattus norv
24	489.5	5.5	2486	4	O75179	O75179 homo sapien
25	488.5	5.5	1486	4	Q8TEF1	Q8tef1 homo sapien
26	486.5	5.5	1009	5	Q8SWY2	Q8swy2 drosophila
27	482.5	5.4	4001	5	Q8WRQ7	Q8wrq7 drosophila
28	481.5	5.4	833	4	Q96I86	Q96i86 homo sapien
29	476.5	5.4	1549	5	Q9V4B1	Q9v4b1 drosophila
30	475.5	5.4	1549	5	Q24241	Q24241 drosophila
31	474.5	5.3	1599	11	Q99NH0	Q99nh0 mus musculu
32	474	5.3	1762	11	Q88521	Q88521 rattus norv
33	473	5.3	2542	4	Q8IWZ3	Q8iwz3 homo sapien
34	473	5.3	2617	4	Q8IWZ2	Q8iwz2 homo sapien
35	466.5	5.3	1188	4	Q9H2B8	Q9h2b8 homo sapien
36	463	5.2	1282	5	Q8T4F8	Q8t4f8 drosophila
37	463	5.2	2119	5	Q9VAU5	Q9vau5 drosophila
38	463	5.2	2119	5	Q8T9G9	Q8t9g9 drosophila
39	453	5.1	638	4	Q8IZ72	Q8iz72 homo sapien
40	451	5.1	1867	5	Q17486	Q17486 caenorhabdi
41	447.5	5.0	1815	5	Q17488	Q17488 caenorhabdi
42	447	5.0	463	4	Q9H6J9	Q9h6j9 homo sapien
43	447	5.0	743	4	Q8TB46	Q8tb46 homo sapien
44	444.5	5.0	919	4	Q8N8A2	Q8n8a2 homo sapien
45	442.5	5.0	1809	5	Q17487	Q17487 caenorhabdi

ALIGNMENTS

RESULT 1
Q9UJH0
ID Q9UJH0 PRELIMINARY; PRT; 1777 AA.
AC Q9UJH0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1250 (Fragment).
GN KIAA1250.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20038619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirosawa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033077; BAA86564.2; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 12.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR PROSITE; PS50088; ANK_REPEAT; 19.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; ANK repeat; Repeat.
FT NON_TER 1 1
SQ SEQUENCE 1777 AA; 197209 MW; B6505923FB45F143 CRC64;

Query Match 99.4%; Score 8860; DB 4; Length 1777;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 0; Indels 58; Gaps 2;

Qy 1 MSVLISQSVINYVERENIPVALKALLEKFDVDERNECGQTPLMIAARQGNLEIVKELIKN 60



RA Mitchell M.; Scnlavo G.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF239046; AAG35185.2; -.
DR HSSP; P80144; 2MY0.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 12.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 12.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1762 AA; 195?.? MW; ? ?????? ??????AB4 CRC64;

Query Match 94.1%; Score ????; DB 1; Length 1762;
Best Local Similarity 91.6%; Pred. ??;
Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;

RESULT 3
Q9ERD4
ID Q9ERD4 PRELIMINARY; PRT; 1715 AA.
AC Q9ERD4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ankyrin repeat and transmembrane-spanning protein.

```
GN   ARMS.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague Dawley;
RX   MEDLINE=22585247; PubMed=12151522;
RA   Kohn H., Boulter J., Weber E.J. [illegible] Diaz M.V.
RT   "An evolutionarily conserved [illegible] protein that is a novel
RT   downstream target of neurotrophin- and ciliary neurotrophin-
RL   J. Neurosci. 21:6 1859-[illegible]
DR   EMBL; AF313464; AAG34167.1; -.
DR   HSSP; P80144; 1MYO.
DR   InterPro; IPR002110; ANK
DR   Pfam; PF00023; ank; 12.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1715 AA;  190532 MW;  [illegible] CA,9AE7D1 B764;

  Query Match             [illegible]%; Score [illegible]; DB 11; Length [illegible];
  Best Local Similarity   [illegible]%; Pred. No. [illegible];
  Matches 1591; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 MSVQDSQSVINYVEEENIPALKATSEYVFLSPNEFSGSLPIDIAAEQNIEIVKSLIKN 60
              ||:  |||| ' |||:   |.|  :|   '| :||  ' : |  :||   |:
Db          1 MSVLDSQSVINYVEEENIPALPATSEYVFLSRNEFSGSLPIDIAAEQNIEIVKSLIKN 60

Qy         61 GANVNLEDLDNWTALIWASKEGHVDVVEFLVSNEIDLNIWTALMWACTYKGHTDVV 120
                |||   |||'  |||   .|    :|  ' :|||  ||'  ||   |||  |.|
Db         61 GANVNLEDLDNWTALIWASKEGHVDVVEFLVSNEIDLNIWTALMWACTYKGHTDVV 120

Qy        121 ELLLSHGANPSVTNQNYVTPLHMAAMNGHVQVVELLLKNGAQVNAQDKFGKTPLDLAAAR 180
              :|||:  :|||.  |||.  :|   :|  :|  |:  |||  |||   |||  |||
Db        121 ELLLSHGANPSVTNLQYVTPLHWAAMMGHVQVVELLLKNGAQVNSQDKFGKTPLDLAAAR 180

Qy        181 KGHLEIVKHLLAMGADVNAQDAMGYTPLHVSVHHLKRNFVMLTLLDKTGNTA 240
              :|||.  |||.  |||   |   :|   ||  ||   ||  |||   |||
Db        181 KGHLEIVKHLLAMGALVDAQDAMGYTPLHVSVHHLKRNFVMLTLLDKTGNTA 240

Qy        241 LMIAARTGHTEIVQELLQAYIVLNVRDKWGRTALHEAVASGYTDIVRL 288
              :|:  .||||  .|||  .|.|                  ||:  '|  |
Db        241 LMIAARTGHTEIVQELLQAYIVLNVRDKWGRTALHEAVASGYTDIVRL 288

Qy        289 QENPTALYWAVEHGATVVDLLAQYCELFTEGLHRDSLENIPRHLRDFWATVSLEYSEFEA 348
              :|:|:  :|||  |||   |.|  :|  :|  |:  |||  |||   |||  |||
Db        289 QENPTALYWAVEHGATVVDLLAQYCELFTEGLHRDSLENIPRHLRDFWATVSLEYSEFEA 348

Qy        349 KVSAVDKKGDTPLHVAIRSRSPTAVSLLLSHGSHQKSIL 420
              |.  ||||.  |||'  |||  .|   ||  |||  ||
Db        349 KVSAVDKKGDTPLHVAIRSRSPTAVSLLLSHGSHQKSIL 420

Qy        421 TQIPGARHLSPTETDGDMLGYDLPYPDHAGMPEDEAPLELTGLDYAQLADSIKSEFLEKL 480
              |.  ||||  ||||.  |||   |   ||  |   |  ||  |||  |||  |||
Db        421 TQIPGARHLSPTETDGDMLGYDLPYPDHAGMPEDEAPLELTGLDYAQLADSIKSEFLEKL 480

Qy        481 EDEMKTPAGQQTIEPLPQPSWLIVPLLIAPTVHLGGIAIGSFLALLYIP 540
              ||  |||.  |||  |||   |   ||  |||  |||  |||  |||
Db        481 EDEMKTPAGQQTEPLPQPSWLIVPLLIAPTVHLGGIAIGSFLALLYIP 540

Qy        541 FIVIYFGGRREGESWNWAWVLSGLAFFCLLFAFPLSFIKALPVRFLF 600
              |||  |||.: ||||  '|  |   |   |  |  ||  ||  |||  |
Db        541 FIVIYFGGRREGESWNWAWALSGIAFFCLLFALPLSFIKALPVRFLF 600

Qy        601 TDYNRLSSVGGETSLAEMIATLSLAVPEFPLACPLFRVPFTEETQSPFKWKKTCLPS 660
              |||  ||||  |||   |'|  .|  :|  |:  |||  ||   |||  |
Db        601 TDYNRLSSVGGETSLAEMIATLSLAVPEFPLACPLFRVPFTEESQSPFKWKKTCLPS 660

Qy        661 FVIFLFIVGCIISGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720
              |||  |||.: ||||  '|  |   |   |  |  ||  ||  |||  |
Db        661 FVIFLFIVGCIIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKPIHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780
              |.:||'  :|||  |||  |||'.|||  :|||||  |||||'||||  :|||  |'|||.
Db        721 NSQRKPIHSAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDAC 780

Qy        781 EQDKVLVMLGTVRVLFSKGPPIAIFASQPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              |.:||'  :|||  |||  |||'.|||  :|||||  |||||'||||  :|||  |'|||.
Db        781 EQDKVLVMLGTVRVLFSKGPPIAIFASQPHIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Qy        841 LPVFLNFPGLSNARKFLVTSATNGDVPISDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900
              ||  |||   ||  |||.  |||  .|:  ||  |,|  |||  :|||  |||||'||||
Db        841 LPVFLNFPGLSNARKFLVTSATNGDITPSDTTGTQEDTDRRVSQNSLGEMTKLGSKTALN 900

Qy        901 RRDTYRRRQMQRTITRQMSFDLTKCLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960
              |||| ||  |||  |||  |||  |||||'||||:|||||:|||||||||||||||||||
Db        901 RRDTYRRRQMQRTITRQMSFDLTKCLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960

Qy        961 FNWDPLAGWINLTEQWPYRTSWLILYLEETEGIPDQMTLKTIYERISKNIPTTKDVEPLL 1020
              |||  :||  |||  |||||:|||||:|||||:|||||:|||||||||||||||||||
Db        961 FNWDPLAGWINLTEQWPYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPITKDVEPLL 1020

Qy       1021 EIDGLIRNPEVELSSRTPVLVARDVKVPLPCTVNLDPKLREIIADVRAAREQISIGGLAY 1080
              |||  |:| '|||  ||||||||  '|||'||||'|||||||||||||  '||||||
Db       1021 EIDGLIRNPEVELSSRTPVLVARDVKTPLPCTVNLDPKLREIIADVRAAREQINIGGLAY 1080

Qy       1081 PPLPLHEGPPRAPSGYSQPPSVCSSTSPNGPFAGGVVSPQPHSSYYSGMTGPQHPFYNRG 1140
              ' |||  ||  |||  ||| ||:||  :||||||||||  ::|||'||||||||||||
Db       1081 PPLPLHEGPPRPPSGYSQPASVCSSASPNGPFPGGVVSPQPHSSYYSGLSGPQHPFYNRA 1140

Qy       1141 SGPAPGFVVLLNSLNVDAVCEKLKQIESLDQSMLPQYCTTIKKANINGRVLAQCNIDELK 1200
              :  | |  :||:: |  '||||:|||| |:|.|:||||,|||||,|:|||||,||
Db       1141 AVPATGSSLLLSSMTVDVVCEKLRQIESGLDQNMMPQYCTTIKKANINGRVLSQCNIDELK 1200

Qy       1201 KEMNMNPYDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHT 1260
              |'| ||  |||. | |  ||: :| |'||||':|| | :||  '.|| ||| '
Db       1201 KEMAMNPYDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPLT 1260

Qy       1261 ELSSQTPYTLNFSPEELNTLGLDEGAPPHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLT 1320
              |||' ||  |||  |||  |||  |||||'||||||||||||||||||||||||||||
Db       1261 ELSSQTPYTLNFSPEELNTLGLDEGAPPHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLT 1320

Qy       1321 DKVQAEYSAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKG 1380
              '|||  |'  |||  |||  |||   ||| :||||'|:|||' ||||' '|||:|
Db       1321 DKVQAEYTAYREYIAQMSQLEGGTGSTTISGRSSPHSTYYIGQSSSGGSIHSTLEQERG 1380

Qy       1381 KLNLEEI  AKGPLYERGDVIDYGNPVSTNEASPLDPITEEDEKSDQSGPKLLPGKKS 1440
              |  |  :|  |||'  |||  |||  |||  |||  |||  |||  |||  |||||
Db       1381 KEGE PLSGRKSPLMKRGDVIDYGNPVSTNEASPLDPITEEDEKSDQSGPKLLPGKKS 1440

Qy       1441 SERSNLFGDLKLKGSGLRYQKLPSDEIESGTEESDNTP-LLKDDKDRKAEGKVERVPKS 1499
              | |  ||  |||  |||  |||  ||  |||  ||  |  |  :|  |||  |||
Db       1441 SERPSLFGDLKLKGSGLRYQKLPSDEIESGTGRVQITPHCSKMIRTKRLKAKQRECASP 1500

Qy       1500 PEHSAEFKTPIKAKEYLSDALLDKKDSSDSGVRSSESSPNHSLHNEVADDSQLEKANLI 1559
              |:||  ||||||:| |||  |||:  ||||:|||||'|||  ,|||||:|||
Db       1501 QEHSAEPRTPIKAKEYLSDALLDKKDSSDSGVRSNESSPNHSLHNEAADDSQLEKANLI 1560

Qy       1560 ELEDLNHPGKRGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYN 1619
              ||:.  :  |||.:|||'  |||  |||  ||||.|||||:|||||||||||:'||||:
Db       1561 ELEDFNHPGKRGMPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYN 1620

Qy       1620 LNRTPSTVTLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHK 1679
              |.||||  ||'|:|, ||||||||  ||||||||||  ||||| ;|| |||| ||
Db       1621 LNRTPSTVTLNNNTAPTNRANQNFDEIPGIRETSQVILRPGPSPNPTAVQNENLKSMAHK 1680

Qy       1680 RSQRSSYTPLSKDPPELHAAASSESTGPGEERESIL 1715
              ||:'||  |||   ||  |||||||| ||||||;
Db       1681 RSQRSSYTPLSKDASELH AASSESTGPGEERESIL 1715

RESULT 4
```

```
Q9UF42
ID   Q9UF42      PRELIMINARY;      PRT;  1031 AA.
AC   Q9UF42;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
GN   DKFZP434F062.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Blum H., Bauersachs S., Mewes H.W., Tiesmeyer J., Wiemann S.;
RL   Submitted (DEC 1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL133620; CAB63746.1; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 10.
DR   PROSITE; PS51297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   1031 AA;  115311 MW;  7DDB6664 713A2F5 CRC64;

  Query Match           58.6%;  Score 2285.5;  DB 4;  Length 1031;
  Best Local Similarity   98.5%;  Pred. No. 0.0;
  Matches 1008;  Conservative    2;  Mismatches    0;  Indels   13;  Gaps    1;

Qy          1 MSVLISQSVINYVEEENIPALAAICATLIVIFVFQNEFQQTPRMIAAFLALEIVKSLIRN 60
              ||||  .|||||  .|||:  | :  .  ||  ||  ||  ||  ||
Db          1 MSVLISQSVINYVEEENIPALAAICATLIVIFVFQNEFQQTPRMIAAFLALEIVKSLIRN 60

Qy         61 GANPNLEDLDNWTALISAAKFGRVEVVQFLIKANIEHKNKWGWTALHWAAYYGRDQVV 120
              ||||  .|||||  .|||:  | :  .  ||  ||  ||  ||  ||  ||  ||
Db         61 GANPNLEDLDNWTALISAAKFGRVEVVQFLIKANIEHKNKWGWTALHWAAYYGRDQVV 120

Qy        121 ELLISHGANPSVTGLQYSVFTIWAAVGRAGIVRLLLQNIAKVNSSDKSGTTPLVWAAR 180
              |.  .|||| .  .||  .|  ||  .  ||  |.| .||  ||
Db        121 ELLLNHGANPSVTGLQYSVFTIWAAVGRAGIVRLLLQNIAKVNSSDKSGTTPLVWAAR 180

Qy        181 KGHLEVVKHLLAYRADVTEPAMDGTTPLHLAVDLQNEDIVSVLLKNGASPNKPNKHGLT 240
              ||||.  .|||  .||  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Db        181 KGHLEVVKHLLAYRADVTEPAMDGTTPLHLAVDLQNEDIVSVLLKNGASPNKPNKHGLT 240

Qy        241 PMDLAKEHGHTEIVQLLDEATTVDSLRMRVAQLQDQVRLMAAMQNISPAMPMPVAPTSS 300
              ||||.  .|||  .||  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Db        241 PMDLAKEHGHTEIVQLLDEATTVDSLRMRVAQLQDQVRLMAAMQNISPAMPMPVAPTSS 300

Qy        301 QDNKTALYWAVEKGNATMVRDILQNNISLINKPDGETPLHLAAKMGNVEVLELLDKDA 360
              ||||| .  |||| .|||.  |||  .  .|.  ||  .  .|  .|  ||  |  |
Db        301 QDNKTALYWAVEKGNATMVRDILQNNISLINKPDGETPLHLAAKMGNVEVLELLDKDA 360

Qy        361 KVSAVDKKGDTPLHIAIRGRSRKIAELLIKNGALVGRLLYRPNKAGETPYDIDCSHQRSIL 420
              ||||||  ||||  .|||  .|.|  .  .|||  |  |||  ||  ||
Db        361 KVSAVDKKGDTPLHIAIRGRSRKIAELLIKNGALVGRLLYRPNKAGETPYDIDCSHQRSIL 420

Qy        421 TQIFGARHLSPTETDGDMLGYDLYSGALADIDFMQPPLYDLYAQWSGKSFILKKL 480
              ||||  ||||  .|||  .|.|  .  .|||  |  |||  ||  ||  ||  ||
Db        421 TQIFGARHLSPTETDGDMLGYDLYSGALADIDFMQPPLYDLYAQWSGKSFILKKL 480

Qy        481 EDEMKTFAGQQIEPLFQFSWLIVFLILLAVSH HLFAFTVHFNLGIAVGLSFLALLYIF 540
              ||||| | ||||  ||||  .|||  ||  ||  ||  ||  |  |||  ||
Db        481 EDEMKTFAGQQIEPLFQFSWLIVFLILLAVSH HLFAFTVHFNLGIAVGLSFLALLYIF 540

Qy        541 FIVIYFGGRREGESWNWAWVLSTRLARHIGYLFLLKLMFVNFIELPEQIIKALPVPFIF 600
              |.||||.| ||| .||||  |||  ||  ||  ||  ||  |  |||  |
Db        541 FIVIYFGGRREGESWNWAWVLSTRLARHIGYLFLLKLMFVNFIELPEQIIKALPVPFIF 600

Qy        601 TDYNRLSSVGGETSLAEMIATLSDADFRFRGIATRLFRVFKTEDTQGKKKWKKTCCLFS 660
              |.||||.| ||| .||||  |||  ||  ||  ||  ||  |  |||  |
Db        601 TDYNRLSSVGGETSLAEMIATLSDADFRFRGFLATRLFRVFKTEDTQGKKKWKKTCCLFS 660

Qy        661 FVIFLFIFGIIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720
              ||. |||  ||  |||  . |||  .||||  ||||  |  ||||  |||||
Db        661 FVIFLFIFGIIISGITLLAIFRVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQVLDSLL 720

Qy        721 NSQRKSLRNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIDGLDAC 780
              ||  |||  ||||  ||  |||  .  ||||  .|||||||||||  |||||
Db        721 NSQRKSLRNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNQTRLVVIDGLDAC 780

Qy        781 EQDKVLQMLLTVRVLFSKGPFIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
              .|||  |.|  |||  |||  ||||  ||||  |||||||||||||||||||||||||
Db        781 EQDKVLQMLLTVRVLFSKGPFIAIFASDPHIIKAINQNLNSVLRDSNINGHDYMRNIVH 840

Qy        841 LPVPINSSGLSNARKFLVTSATNGDVPQSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900
              |||  .|||  ||||.|||  ||||  |||||||||||||||||||  ||||  ||||
Db        841 LPVPINSSGLSNARKFLVTSATNGDVPQSDTTGIQEDADRRVSQNSLGEMTKLGSKTALN 900

Qy        901 RRDTYRRPQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960
              |  |||  |||  |||  |||  ||||  ||||||||||||  ||||  ||||  ||
Db        901 RRDTYRRPQMQRTITRQMSFDLTKLLVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 960

Qy        961 FNWDRLAFWINLTEQWPYPTSWLILYLEETEGIPDQMTLKTIYER---  ----  -IS 1007
              ||||  |||  ||||  |||  ||||  ||||||||||||  ||  ||
Db        961 FNWDRLAFWINLTEQWPYPTSWLILYLEETEGIPDQMTLKTIYERCCGADSTDRDRIGIS 1020

Qy       1008 KNI 1010
              |||
Db       1021 KSV 1023


RESULT 5
Q9H9E4
ID   Q9H9E4      PRELIMINARY;      PRT;  543 AA.
AC   Q9H9E4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ12811.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Ono T., Kobayashi K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsuzawa H., Ishii S., Fuwai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022873; BAB14285.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   543 AA;  59976 MW;  F5D643D5A20C641D CRC64;

  Query Match           31.4%;  Score 2791;  DB 4;  Length 543;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-176;
  Matches  541;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1173 MLPQYCTTIKKANINGRVLAQCNIDELKKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDP 1232
              ||  |||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLPQYCTTIKKANINGRVLAQCNIDELKKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDP 60

Qy       1233 RFLSESSSGPAPHGEPARRASHNELPHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 1292
              ||.|||   ||||  |||  ||||||||||||||||||||||||||||||||||||||||
Db         61 RFLSESSSGPAPHGEPARRASHNELPHTELSSQTPYTLNFSFEELNTLGLDEGAPRHSNL 120

Qy       1293 SWQSQTRPTPSLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISG 1352
              |||  |||  |||  ||||.|||||  ||||.|||||||||||  ||||||  .||||
Db         121 SWQSQTRPTPSLSSLNSQDSSIEISKLTDKVQAEYRDAYREYIAQMSQLEGGPGSTTISG 180

Qy       1353 RSSPHSTSTWQQSSSGGSIHSNLEQSKGKDSEPKPDDGRKSFLMKRGDVIDYSSSGVSTN 1412
```

```
Db        181 PSSPHGTYYMGQGSSQGGSTHGNDAEPGKIEEKPDQGRPGRLMKPGCIDYGSSWRIN 240

Qy       1413 DASPLDPITEEDEKSLSNPIIPPPAPPPPLPQTQLMPPVGLPAYLPLPSDPLPIST 1472

Db        241 DASPLDPITEEDEKSLSNPIIPPPAPPPPLPQTQLMPPVGLPAYLPLPSDPLPIST 300

Qy       1473 EDSINTPLLKQGAEPPAPKKPRVTLDLFSDEEIIPTPTPAPEYLNTA-PPKLLDNGV 1532

Db        301 EDSLNTPLLKQGVEPPAPKKPRVTLDLFSDEEIIPTPTPAPEYLNTAAPPKLLDNGV 360

Qy       1533 RSGRSGPNHSLHDAVALGPPPGLHPAGPSSLSGSGLSQLGIAPVHLQE 1582

Db        361 RSSRSGPNRSLHDSVATGPPPRLHPAGPSSLSGSGLSQLGIAPVHLSE 410

Qy       1583 DKKSTPPCSLIAASPPEGWPAVQPAYNLNPIKTVTLNNNPAPANPAN-NPDEMRTRET 1642

Db        411 DKKSTSPCSLIAASPPEGWPAVQPAYNLNPIKTVTLNNNPAPANPANANPDEMRTRET 470

Qy       1643 SQVTLPPSANPNETTTQALNTEKTPAPSQTVPTPLPATTSLPSTTAAA-EGTGPAERRE 1701

Db        471 SQVTLPPSANPNETTTQALNTEKTPAPSQTVPTPLPATTSLPSTTAAAAEGTGPAERRE 530

Qy       1702 SIL 1704

Db        531 SIL 533

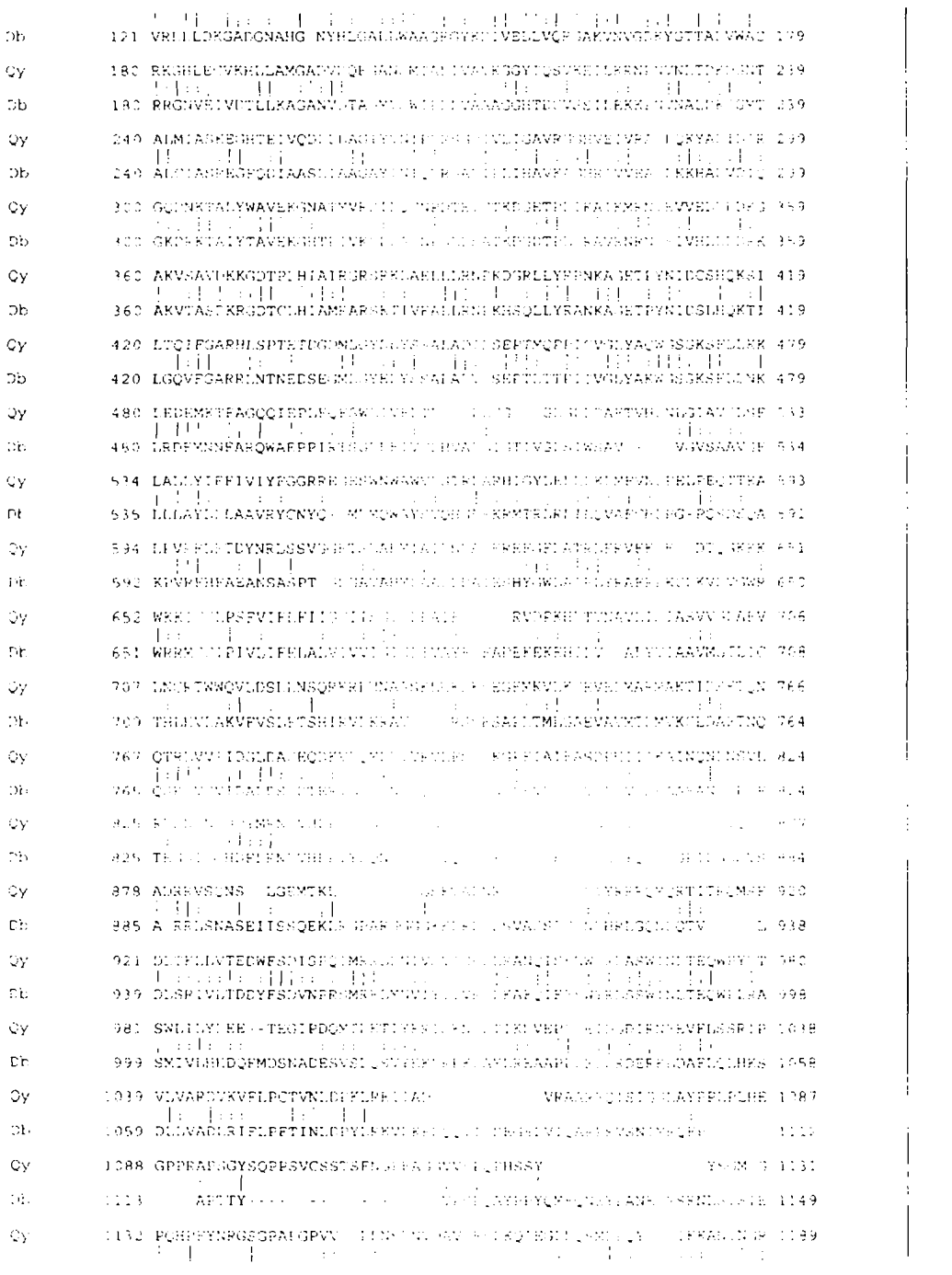

RESULT 4
Q9W2I0
ID   Q9W2I0      PRELIMINARY;      PRT;   534 AA.
AC   Q9W2I0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 21, Last annotation update)
DE   CG30187 protein.
GN   CG30187 OR Dmel\CG30187 OR CG30187.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20[illegible]; PubMed=[illegible];
RA   Adams M.D., [illegible];
RA   Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G.,
RA   Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A.,
RA   Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G.,
RA   Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C.,
RA   Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H.,
RA   Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Bangon C., An H., Baldwin D., Bangon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Krause D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunez J.,
RA   Pacleb J., Parsons V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Berman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Shue B., Smutniak F., Whitfield E.,
RA   Ashburner M., Hartl D.L., Hart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN
RP   SEQUENCE FROM N.A.
RA   Misra S.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003745; AAF46711.2; -.
DR   FlyBase; FBgn0030987; CG30187.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001487; BTB_LysR.
DR   Pfam; PF00023; ank; 11.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 12.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00944; BCH_LYSR_FAMILY; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1476 AA;  162594 MW;  15F8376E3F27B843 CRC64;

  Query Match             26.8%;  Score 2379.5;  DB 5;  Length 1476;
  Best Local Similarity   37.6%;  Pred. No. 2.3e-148;
  Matches  663;  Conservative  249;  Mismatches  490;  Indels  179;  Gaps   39;

Qy          1 MSVLLQGVINYVREENIPALKALLE-KCKDVDERNECGQTPLMIAAEQGNLEIVKELIK 59
              |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |  |  |  |  |
Db          1 MGSIGRAIQQYIDNNDISGLRAILDSRHLTIDDRDENATTVLMVVAGRGLTAFVREFLA 60

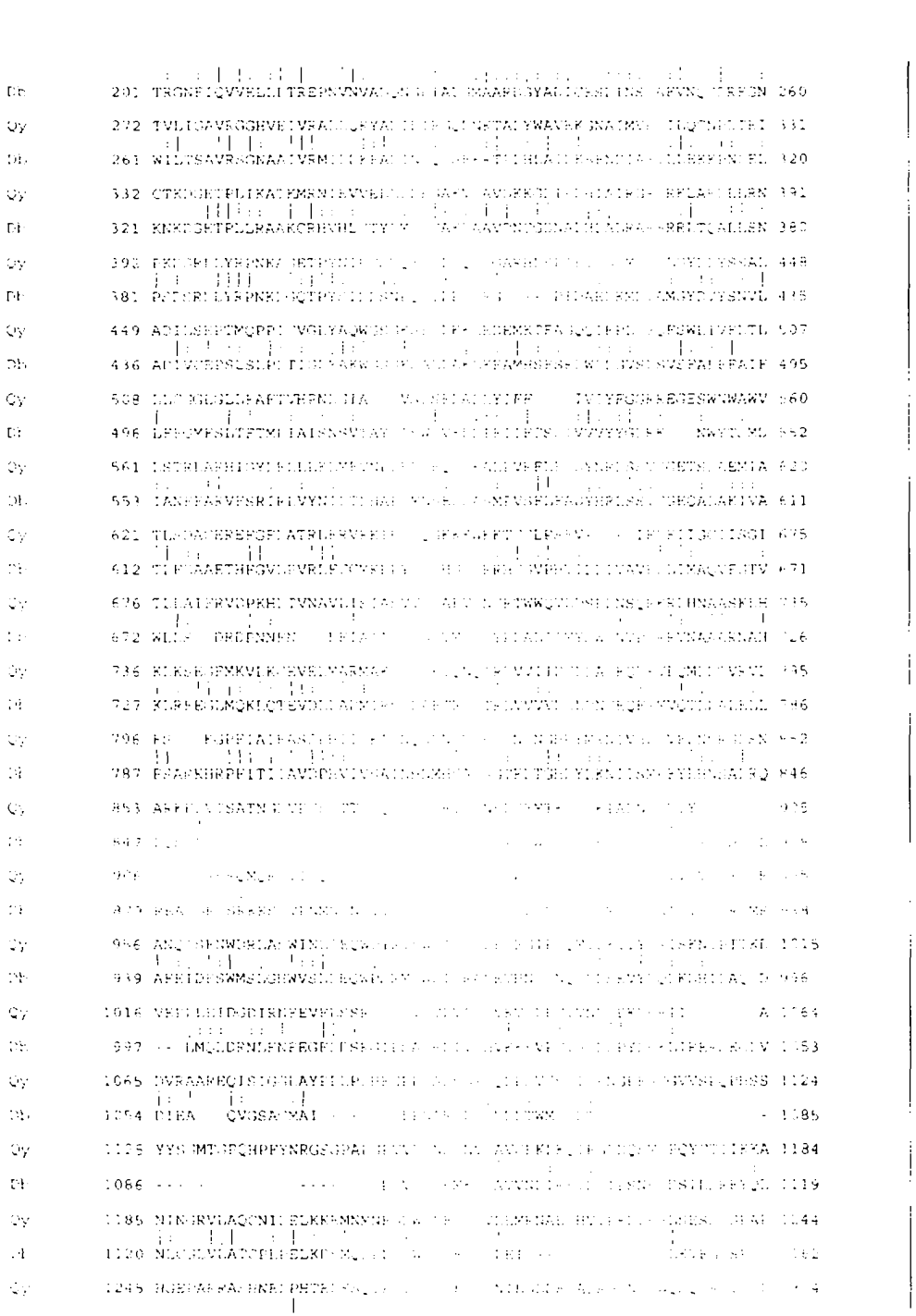
Qy         60 NGANVNCFDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDV 119
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db         61 RGADVQAQDKLDNWTALLCASRNGHLDVVQLLLDHGAEVEHRDMGGWTSLMWAAYRGHTEI 120

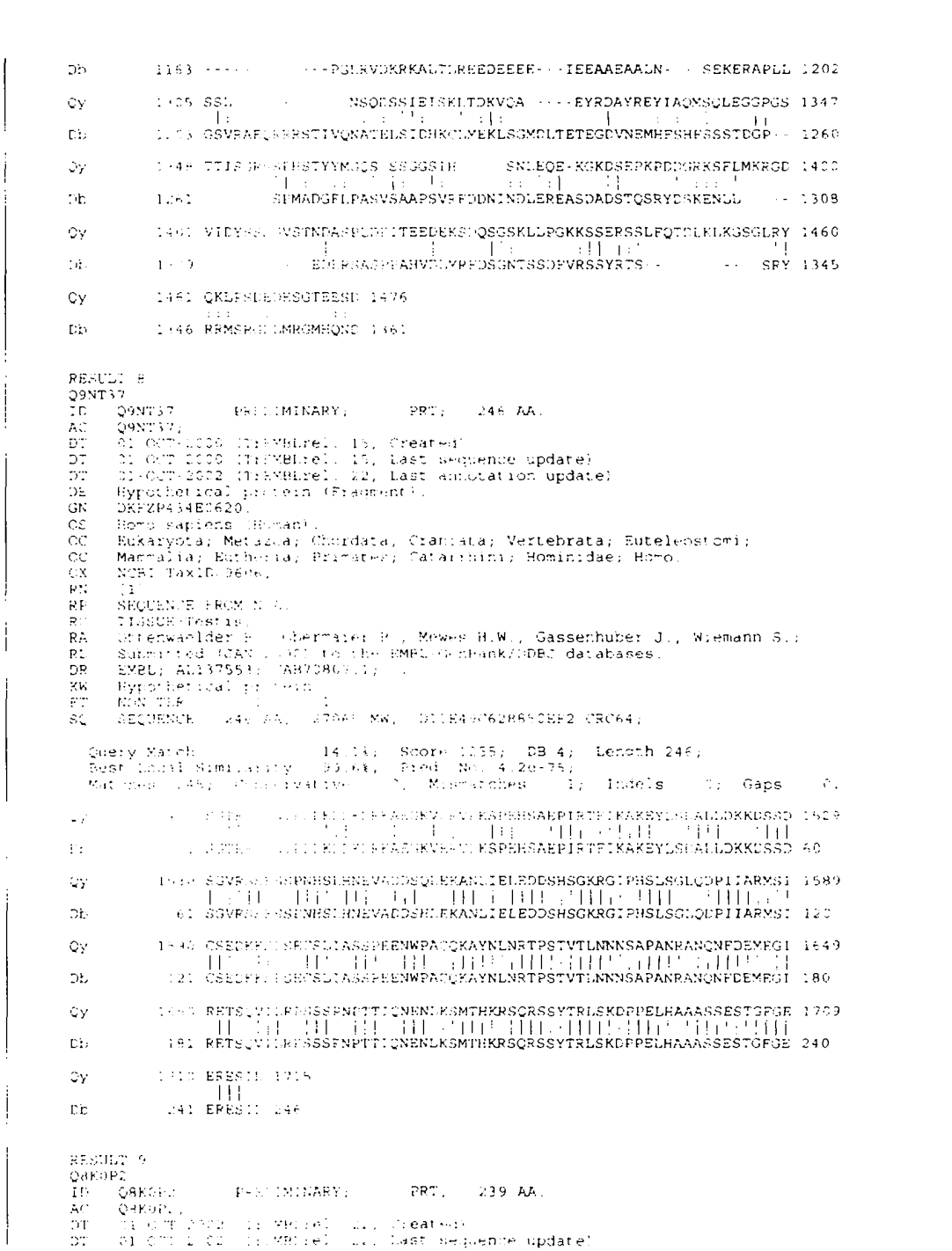
Qy        120 VELLLKHGSVTGLQYGVYPITWAARRGHADIVHLLLQNGAKVNCSDKYGTTPLYWAA 176
```



RESULT 7
Q20109
ID Q20109 PRELIMINARY; PRT; 1398 AA.
AC Q20109;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE F36H1.2 protein.
GN F36H1.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw C.K.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069612; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z69760; CAA93996.1; -.
DR HSSP; P00420; 1AR7.
DR WormPep; F36H1.2; CE05814.
DR InterPro; IPR...; ANK.
DR InterPro; IPR...; TRNA_synt_1.
DR Pfam; PF...; ANK; 10.
DR SMART; SM00248; ANK; 9.
DR PROSITE; PS50887; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1398 AA; 156568 MW; D5080A9E89AB1BEC CRC64;

Query Match 20.6%; Score 1832.5; DB 5; Length 1398;
Best Local Similarity 30.9%; Pred. No. 4.1e-112;
Matches 468; Conservative 280; Mismatches 483; Indels 285; Gaps 40;





```
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Similar to RIKEN cDNA 2110039L19 gene (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN 2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC030876; AAH30876.1;
FT   NON_TER       1       1
SQ   SEQUENCE   249 AA;  26184 MW;  7D12D9DCF9 CRC64;

  Query Match             12.6%;  Score 113.5;  DB 11;  Length 249;
  Best Local Similarity   89.6%;  Pred. No. 3.7e-66;
  Matches  215;  Conservative   12;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       1476 DNTPLLKDDKDKKAEGKAERVAKSPEHSVPIPTFIKAKEYLSDALLDKFDSSDSGVRSS 1535
             ||||   |||:: ||| : | ||: : | |   |||| ||: ||   ||.:||.::
Db          1 DNTPLLKDDKDKKAEGKAERVAKSPEHSVPIPTFIKAKEYLSDALLDKFDSSDSGVRSN 60

Qy       1536 ESSPNHSDHNEVADDSQIEKANIIELEDGHHPPGMPHSIGDLQDPVIARMSICSDDKK 1595
             ||   |:|:: |||   ||| |'.. ':|.:::|  ||   |.  ||| :||
Db         61 ESSPNHGDHNEAADDSQIEKANIIELEDGHHPPGMPHSIGDLQDPVIARMSICSDDKK 120

Qy       1596 SPSPCSLIASSPEENWFAQKAYNLNRTPSTVTLNNNSAPANRANQNPDAMEGIRETSQV 1655
             |||:: ||| |':| ::||' : | : :|: |:| :|| ||| :':  |||
Db        121 SPSPGSLIASSPEESWPSTQKAYNLNRTPSTVTLNNNTAPTNRANQNPDAMEGVRETSQV 180

Qy       1656 ILRPSSSPNPTTIQNENLKSMHPPSLGSYTDLSKDPPEIHAAASSEMTSGEERESIL 1715
             |||:  |||: :|||: '' :  : :  ||:  |  || :||: :||.:||
Db        181 ILRPGTSPNPTAVQNENLKSMAHPPSLGSYTDLSKDADLDM AASSGTTFGEERESIL 239


RESULT 10
Q9H889
ID   Q9H889      PRELIMINARY;      PRT;   129 AA.
AC   Q9H889;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ13264
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RA   Isogai T., Ota T., Hayashi K., Otsuki T., Nunoya T., Ono T., Sugiyama T.,
RA   Nishikawa T., Nagai K., Sugano S., Isogai T., Suzuki Y., Hata H.,
RA   Tanase T., Nomura Y., Togiya S., Komaki F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Sait O., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Muramo Y., Ishima Y.,
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK023926; BAB14278.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   129 AA;  14744 MW;  4BF8F1A8...348BF CRC64;

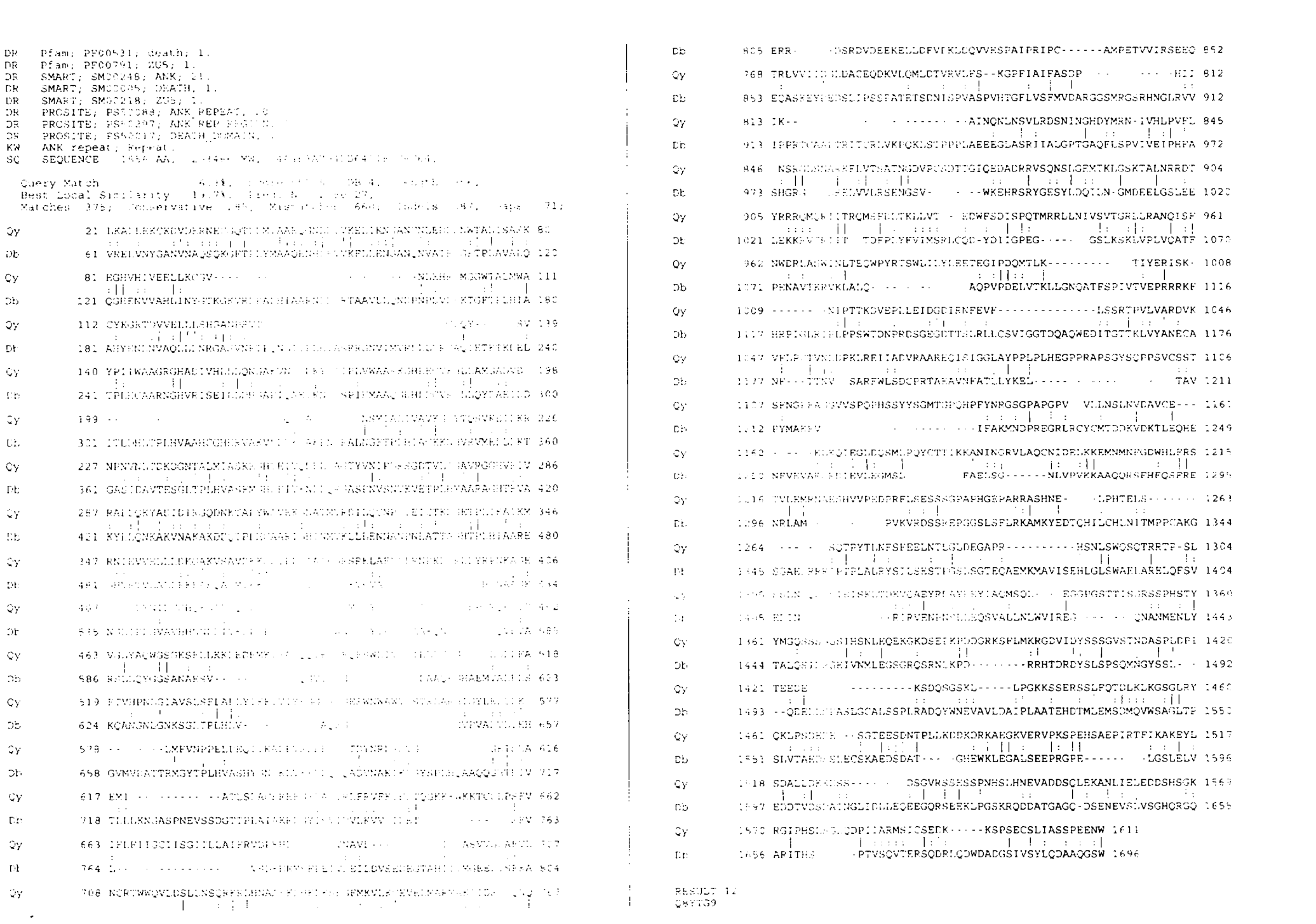
  Query Match              7.5%;  Score 67.5;  DB 4;  Length 129;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-36;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1587 MSICSEDKKSPSECSLIASSPEENWFAQKAYNLNRTPSTVTLNNNSAPANRANQNPDAM 1646
             | | ||||||||||| |||| ||| ||| |||| ||||||||| ||| ||| ||
Db          1 MSICSEDKKSPSECSLIASSPEENWFAQKAYNLNRTPSTVTLNNNSAPANRANQNPDAM 60

Qy       1647 EGIRETSQVILRPSSSPNPTTIQNENLKSMHPPSLGSYTDLSKDPPEIHAAASSEMTSG 1706
             |.. ||||:: |||| '||| : ||| ||| |:| ||| :|| :|| ||| ||.  :|
Db         61 EGIRETSQVILRPSSSPNPTTIQNENLKSMHPPSLGSYTDLSKDPPEIHAAASSEMTSG 120

Qy       1707 PGEERESIL 1715
              |||: |||
Db         121 PGEERESIL 129


RESULT 11
Q99407
ID   Q99407      PRELIMINARY;      PRT;  1856 AA.
AC   Q99407;
DT   01-MAY-1997 (TrEMBLrel. 01, Created)
DT   01-MAY-1997 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 21, Last annotation update)
DE   Ankyrin.
GN   ANK1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97382244; PubMed=9235914;
RA   Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;
RT   "Structure and organization of the human ankyrin-1 gene. Basis for
RT   complexity of pre-mRNA processing.";
RL   J. Biol. Chem. 272:19220-19228(1997).
DR   EMBL; U50133; AAB47805.1; -.
DR   EMBL; U50092; AAB47805.1; JOINED.
DR   EMBL; U50093; AAB47805.1; JOINED.
DR   EMBL; U50094; AAB47805.1; JOINED.
DR   EMBL; U50095; AAB47805.1; JOINED.
DR   EMBL; U50096; AAB47805.1; JOINED.
DR   EMBL; U50097; AAB47805.1; JOINED.
DR   EMBL; U50098; AAB47805.1; JOINED.
DR   EMBL; U50099; AAB47805.1; JOINED.
DR   EMBL; U50100; AAB47805.1; JOINED.
DR   EMBL; U50101; AAB47805.1; JOINED.
DR   EMBL; U50102; AAB47805.1; JOINED.
DR   EMBL; U50103; AAB47805.1; JOINED.
DR   EMBL; U50104; AAB47805.1; JOINED.
DR   EMBL; U50105; AAB47805.1; JOINED.
DR   EMBL; U50106; AAB47805.1; JOINED.
DR   EMBL; U50107; AAB47805.1; JOINED.
DR   EMBL; U50108; AAB47805.1; JOINED.
DR   EMBL; U50109; AAB47805.1; JOINED.
DR   EMBL; U50110; AAB47805.1; JOINED.
DR   EMBL; U50111; AAB47805.1; JOINED.
DR   EMBL; U50112; AAB47805.1; JOINED.
DR   EMBL; U50113; AAB47805.1; JOINED.
DR   EMBL; U50114; AAB47805.1; JOINED.
DR   EMBL; U50115; AAB47805.1; JOINED.
DR   EMBL; U50116; AAB47805.1; JOINED.
DR   EMBL; U50117; AAB47805.1; JOINED.
DR   EMBL; U50118; AAB47805.1; JOINED.
DR   EMBL; U50119; AAB47805.1; JOINED.
DR   EMBL; U50120; AAB47805.1; JOINED.
DR   EMBL; U50121; AAB47805.1; JOINED.
DR   EMBL; U50122; AAB47805.1; JOINED.
DR   EMBL; U50123; AAB47805.1; JOINED.
DR   EMBL; U50124; AAB47805.1; JOINED.
DR   EMBL; U50125; AAB47805.1; JOINED.
DR   EMBL; U50126; AAB47805.1; JOINED.
DR   EMBL; U50127; AAB47805.1; JOINED.
DR   EMBL; U50128; AAB47805.1; JOINED.
DR   EMBL; U50129; AAB47805.1; JOINED.
DR   EMBL; U50130; AAB47805.1; JOINED.
DR   EMBL; U50131; AAB47805.1; JOINED.
DR   EMBL; U50132; AAB47805.1; JOINED.
DR   HSSP; Q00402; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
```


DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 19.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1656 AA;
Query Match
Best Local Similarity
Matches 376;
Qy
Db
RESULT 12
Q8TTG9

```
ID   Q8YTG9      PRELIMINARY;      PRT,   414 AA.
AC   Q8YTG9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein All2748.
GN   ALL2748.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21535285; PubMed=11759494;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasaki S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; AP003590; BAB74447.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   414 AA;  43900 MW;  [illegible] CRC64;

  Query Match             41.1%;  Score 544;  DB 2;  Length 414;
  Best Local Similarity   30.2%;  Pred. No. 1.1e-29;
  Matches  149;  Conservative  144;  Mismatches  186;  Indels   8;  Gaps   6;

Qy          9 VINSVERENIPAIPALLERGYRVDFKDYMGQTQLHIAAALGHIEIVEFLLKNGANVNLQE 67
              ::  :  :| :  ||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db          8 LLKVARSGDTKGIMALLAAGVNVNFQNQFGETALMFAANLGHLEIVRSLLSSGADINIAR 67

Qy         68 DLDNWTALISASKEGHVRIVEVLLKYGADRHMQDWTALYWATYRGHVEIVELLIKHG 127
              :|: :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         68 KPYYPTALMLAASNQVDIVKYLLSNGADVNVQDCAGRTALHFAVGAGYPDIVKLLLSAG 127

Qy        128 ADLNT [illegible] 187
              :  :
Db        128 AIVSI [illegible] 187

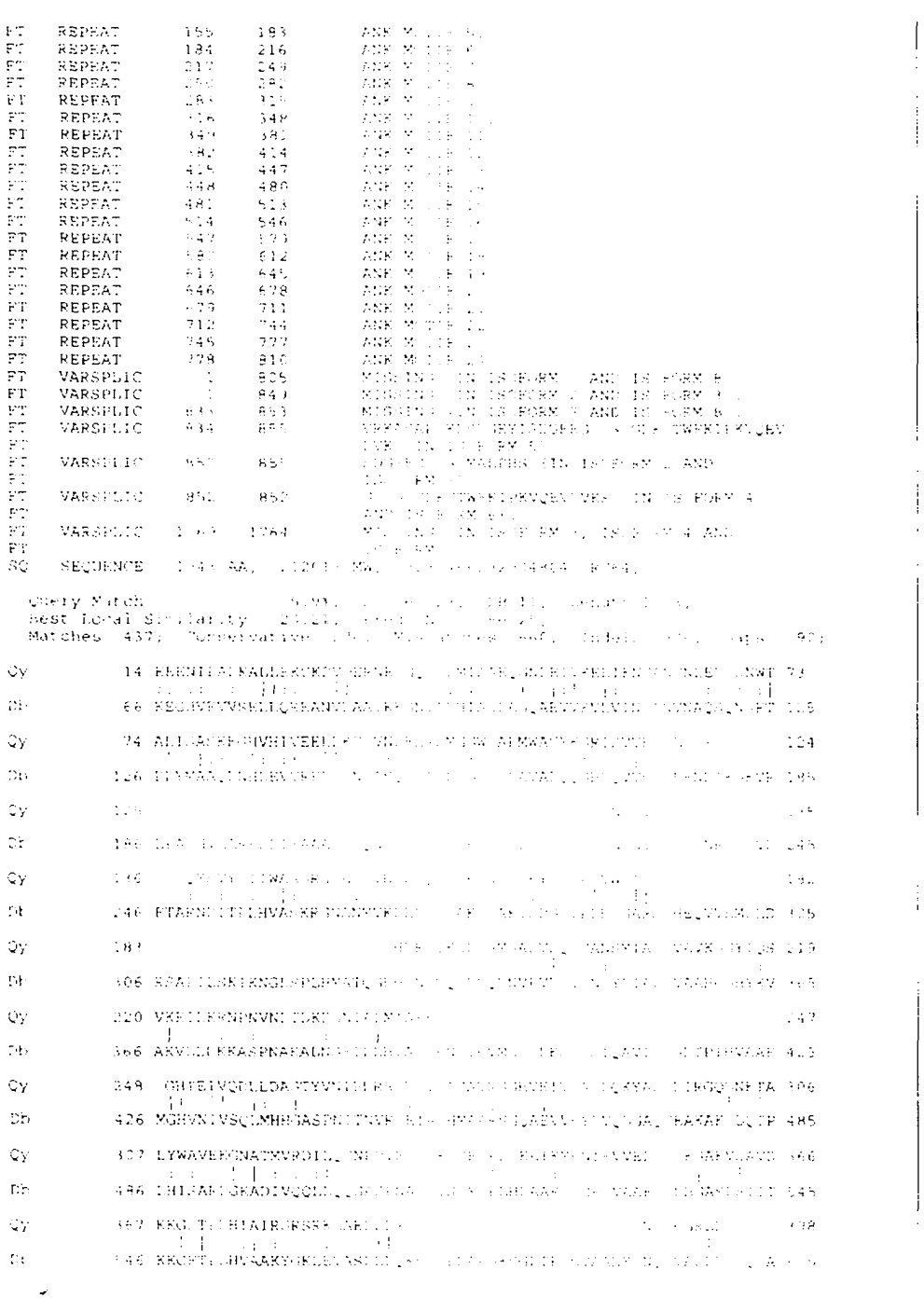
Qy        188 KRLLAMSADVDQRANSMALVLA [illegible] 247
              :  :| :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        187 QVELAAGVDVNEQNQDGSTALAAVAA [illegible] 246

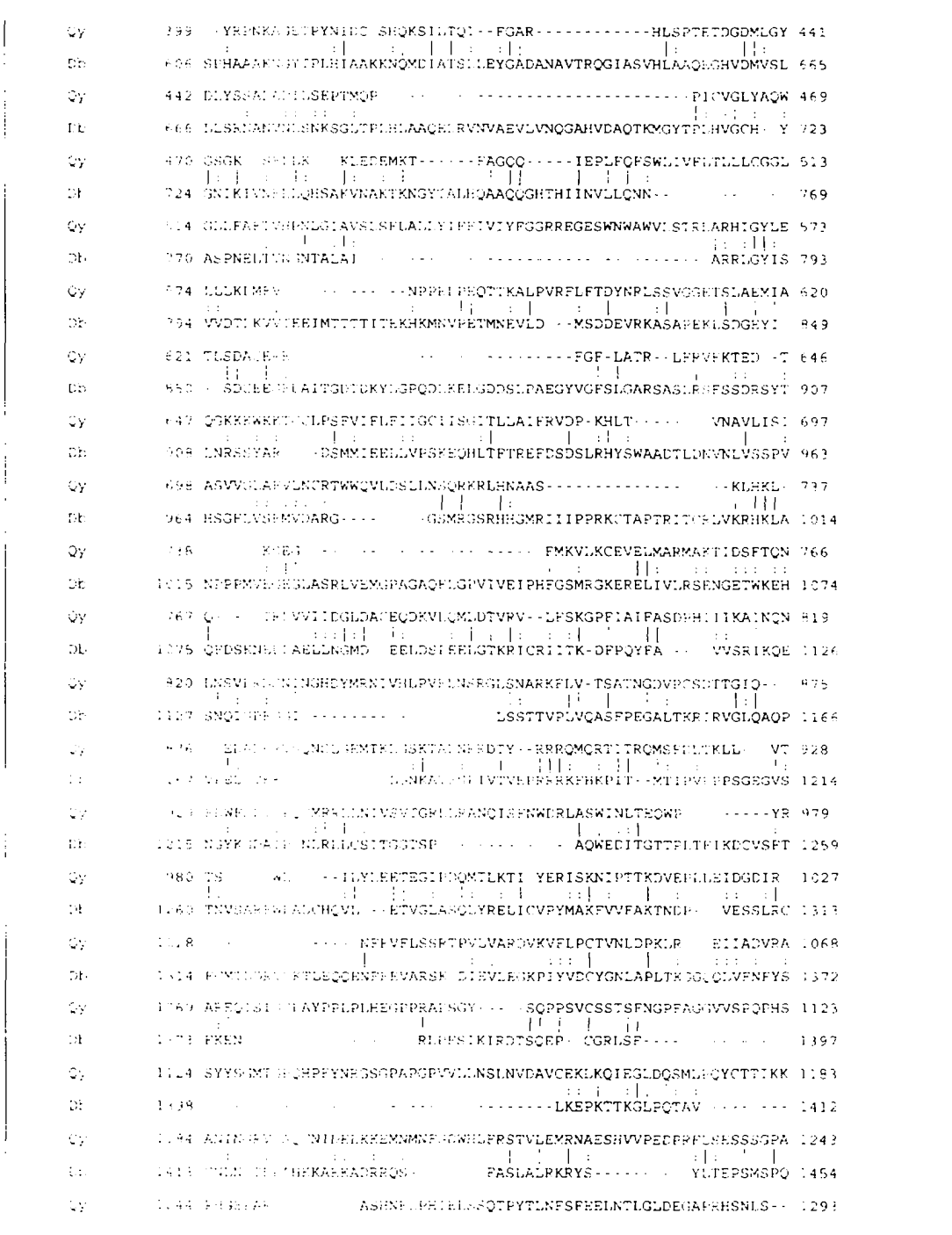
Qy        248 GHCEIVQVLLDAGTYVNIPDRLGRTPVDVAVDNGHEDIAEYLQQAMTA 296
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        247 GYVDVVQVLLNQGANTQIRNKLGRTPLDVAAL [illegible] 304

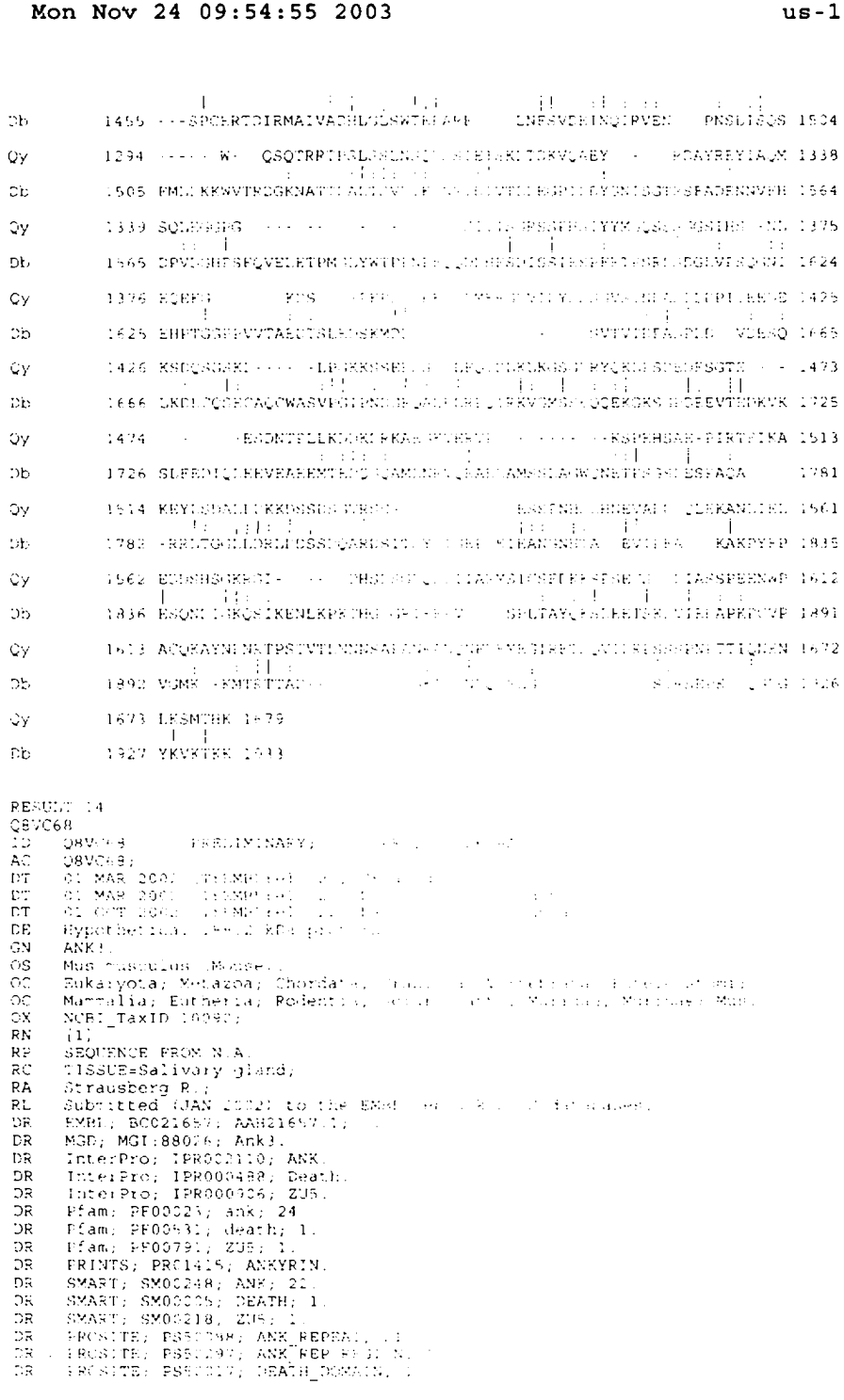
Qy        297 LYWAVEKGNATMVRDILQ [illegible] 356
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        305 ITLAASQGHTATVRILLDYGAKANIPAKDKTPLDLAKEFKHTEIVKLLE 354

Qy        357 KKGDTPLHIAIRGRSRKLASLLIDNPP [illegible] 414
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        355 SVGATALIWAASGGYNKVVQILQE [illegible] 412


RESULT 13
Q61307
ID   Q61307      PRELIMINARY;      PRT,  [illegible] AA.
AC   Q61307; Q61305; Q61306; Q61308, [illegible] Q61310, [illegible] [illegible];
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ankyrin 3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin [illegible]).
GN   ANK3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090.
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 6)
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=95340[illegible]; PubMed=7615634;
RA   Peters L.L., Otto E.M., Lu F.M., Eicher E.M., Higgins A., Yialamas M.,
RA   Turtzo L.C., [illegible] A.J., Lux S.E.;
RT   "Ank3 (epithelial ankyrin), a widely distributed new member of the
RT   ankyrin gene family and the major ankyrin in kidney, is expressed in
RT   alternatively spliced forms, including forms that lack the repeat
RT   domain.";
RL   J. Cell Biol. 130:313-330(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RC   STRAIN=C3H/HeJ; TISSUE=Bone Marrow;
RX   MEDLINE=97213781; PubMed=9060470;
RA   Hoock T.C., Peters L.L., Lux S.E.;
RT   "Isoforms of Ankyrin-3 that lack the NH2-terminal repeats associate
RT   with mouse macrophage lysosomes.";
RL   J. Cell Biol. 136:1059-1070(1997).
CC   -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
CC       SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
CC       BAND 4.2, TO NA K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
CC       AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
CC       DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
CC       THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC       THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC       ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1 (SHOWN HERE), 2, 3,
CC       4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC       EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC       LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC       SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC   -!- SIMILARITY: CONTAINS 24 ANK REPEATS.
DR   EMBL; L40631; AAB01603.1; -.
DR   EMBL; L40644; AAB01605.1; -.
DR   EMBL; U69274; AAB58481.1; -.
DR   EMBL; L47[illegible]; AAB[illegible]; -.
DR   EMBL; L40[illegible]; AAB[illegible]; -.
DR   EMBL; L40[illegible]; AAB[illegible]; -.
DR   EMBL; L40[illegible]; AAB[illegible]; -.
DR   EMBL; U69274; [illegible]; -.
DR   HSSP; P55273; 1svf.
DR   MGD; MGI:88028; Ank3.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000306; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Multigene family
FT   DOMAIN        1    826       ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT   DOMAIN       23    910       24 X ANK MOTIF REPEATS.
FT   DOMAIN     [illegible]  1437  SPECTRIN BINDING DOMAIN.
FT   DOMAIN     [illegible]  1943  REGULATORY DOMAIN
FT                                 (REGULATES BINDING OF ANKYRIN TO SPECTRIN
FT                                 AND THE BAND 3 PROTEIN).
FT   REPEAT       23    55        ANK MOTIF 1.
FT   REPEAT       56    88        ANK MOTIF 2.
FT   REPEAT       89    121       ANK MOTIF 3.
FT   REPEAT      122    154       ANK MOTIF 4.
```



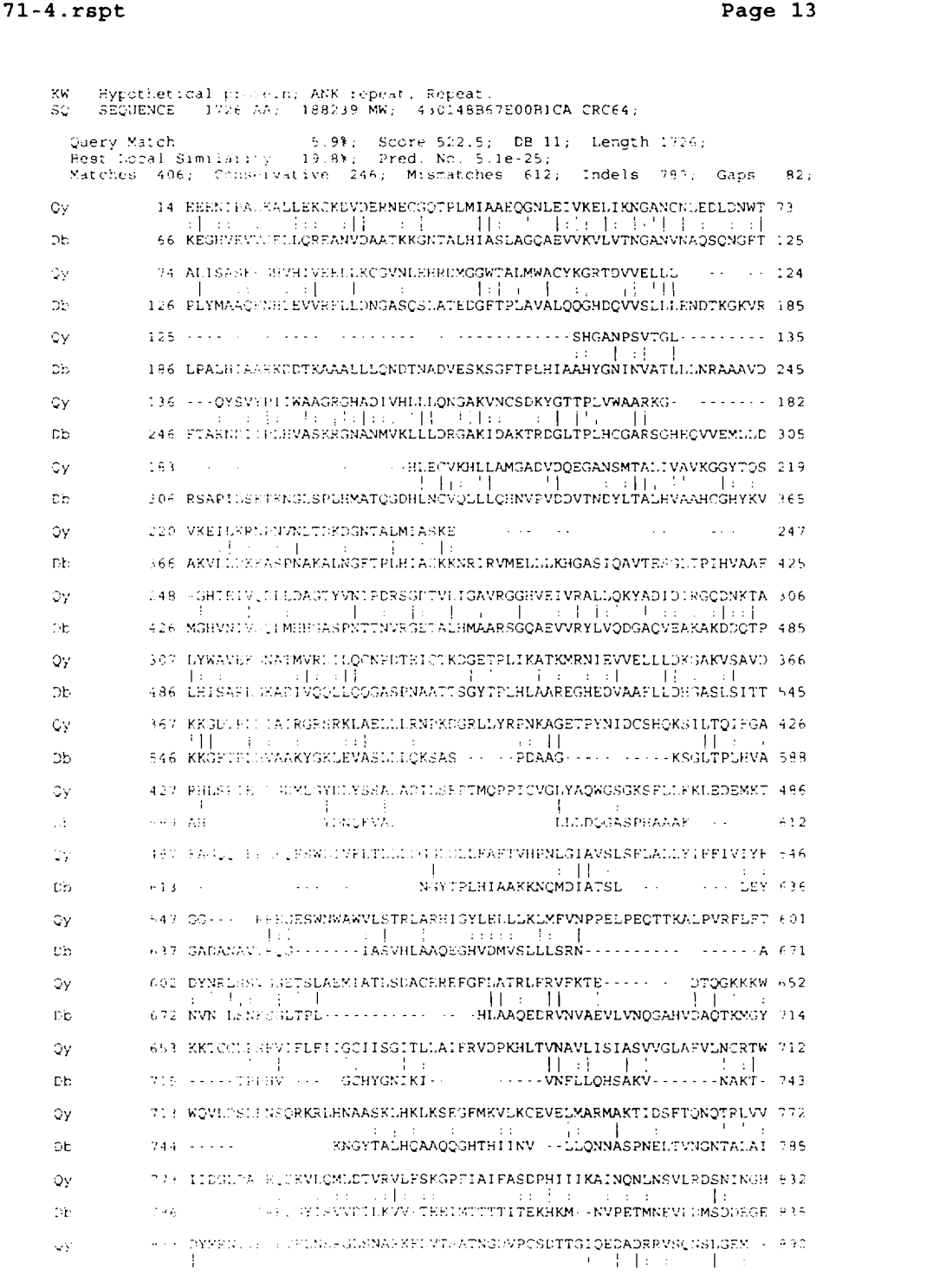


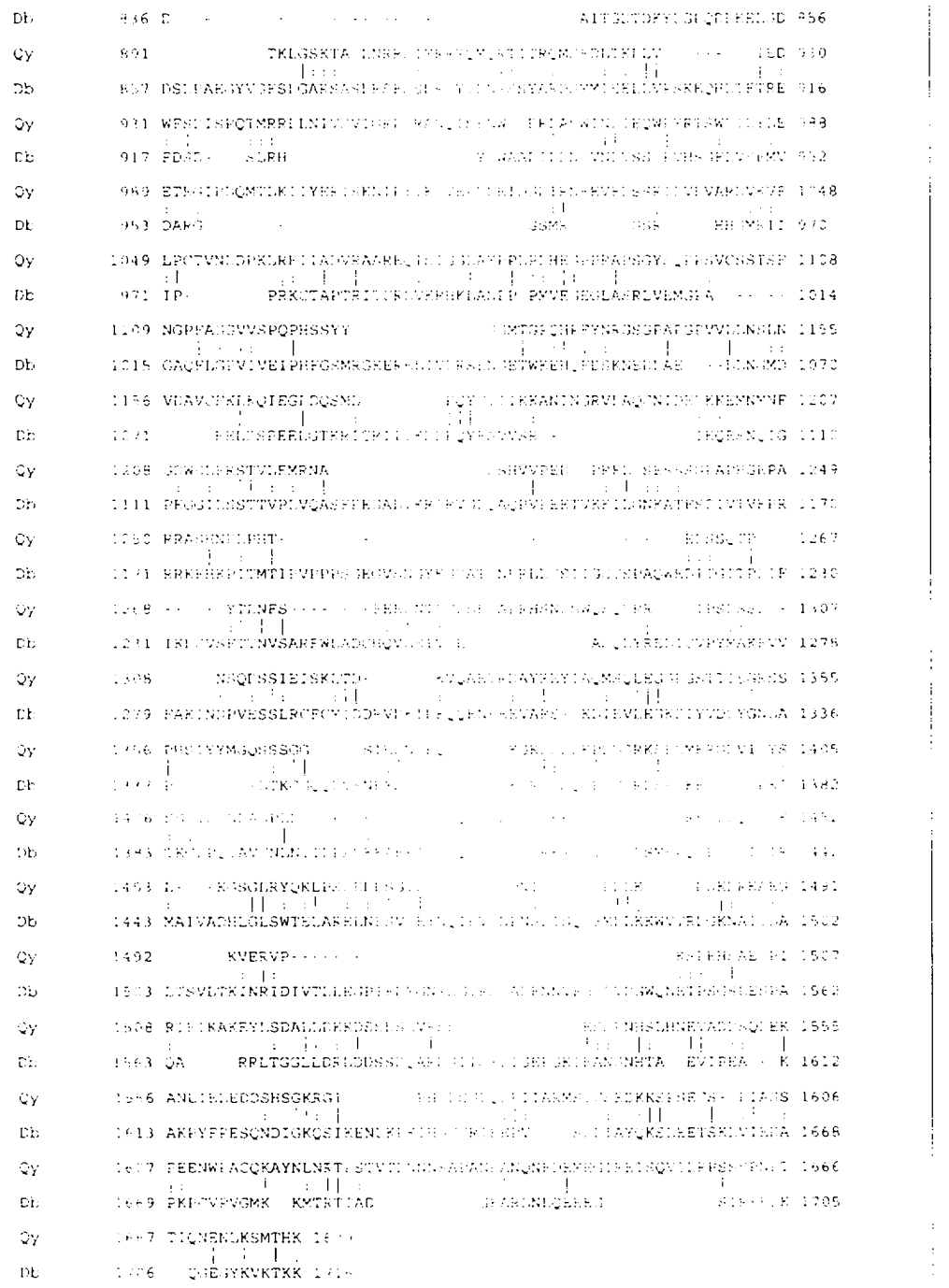


RESULT 14
Q8VC68
ID Q8VC68 PRELIMINARY; [illegible]
AC Q8VC68;
DT 01-MAR-2002 (TrEMBLrel. [illegible])
DT 01-MAR-2002 (TrEMBLrel. [illegible])
DT 01-OCT-2002 (TrEMBLrel. [illegible])
DE Hypothetical [illegible] KDEL protein.
GN ANK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata, [illegible] Vertebrata; [illegible] Mammalia;
OC Mammalia; Eutheria; Rodentia, [illegible] Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC021657; AAH21657.1; -.
DR MGD; MGI:88026; Ank1.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 13.
DR PROSITE; PS50297; ANK_REP_REGION; 5.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.

KW Hypothetical protein; ANK repeat; Repeat.
SQ SEQUENCE 1726 AA; 188239 MW; 4301488B67E00B1CA CRC64;

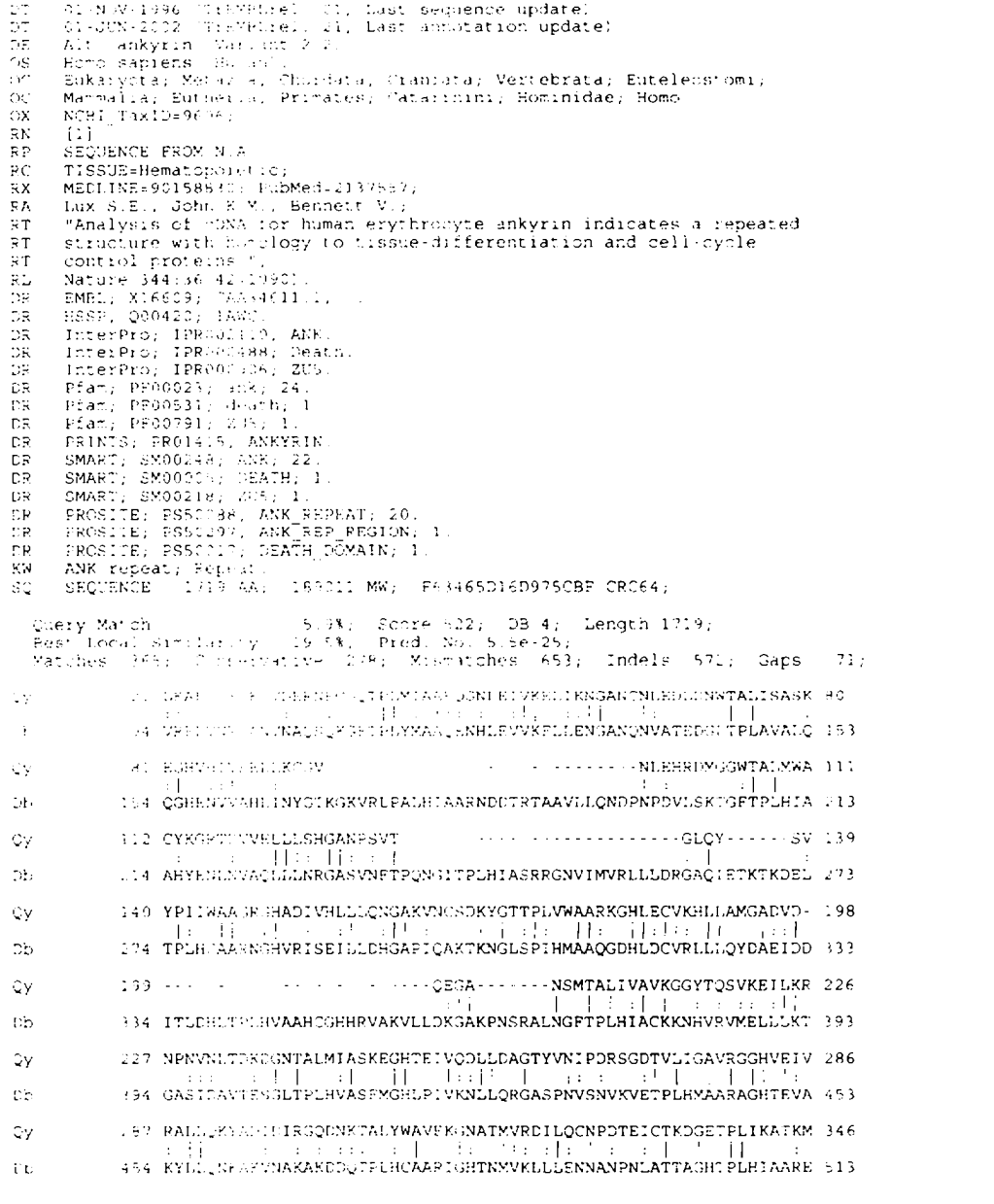
Query Match 5.9%; Score 522.5; DB 11; Length 1726;
Best Local Similarity 19.8%; Pred. No. 5.1e-25;
Matches 406; Conservative 246; Mismatches 612; Indels 793; Gaps 82;

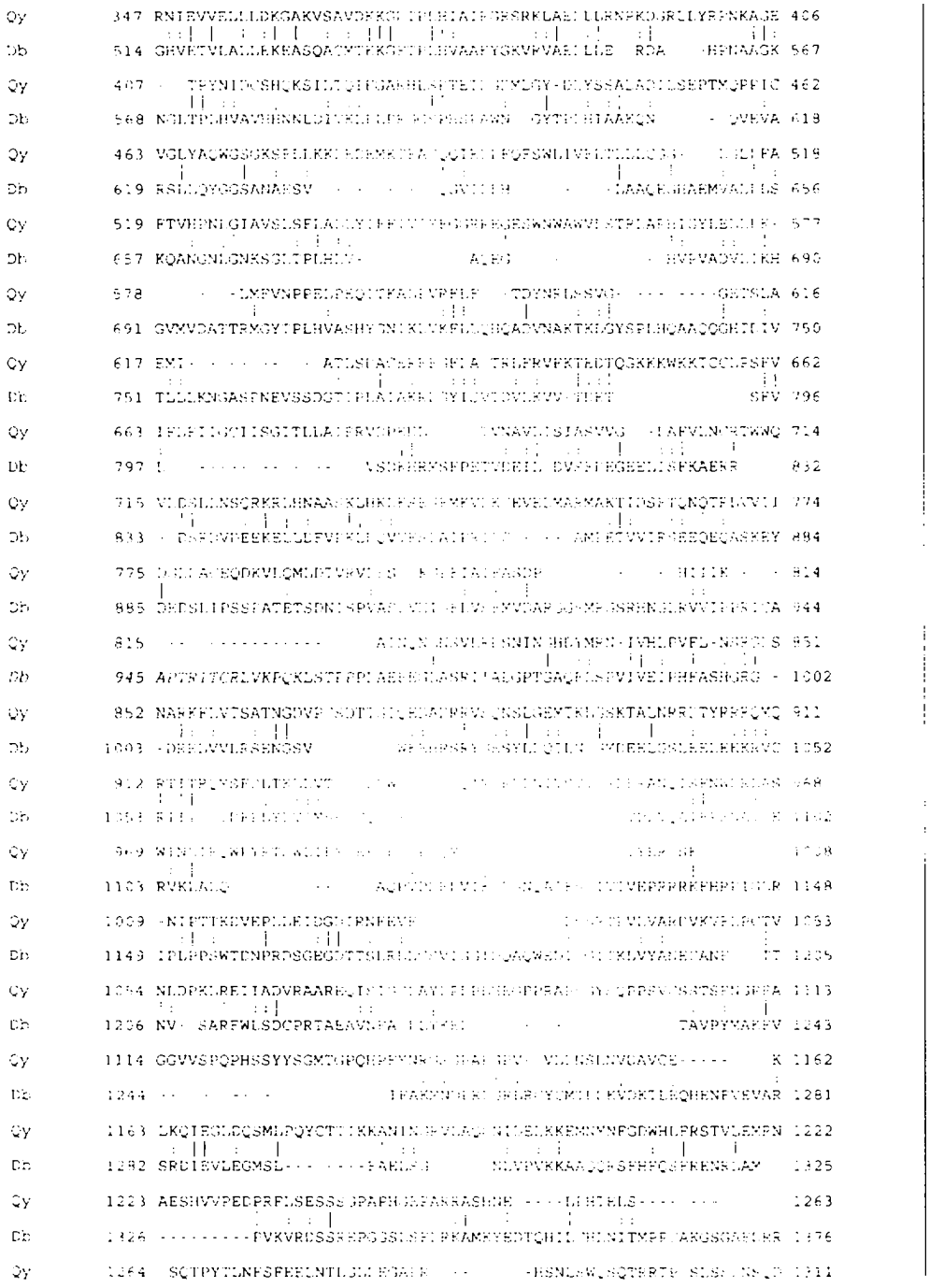


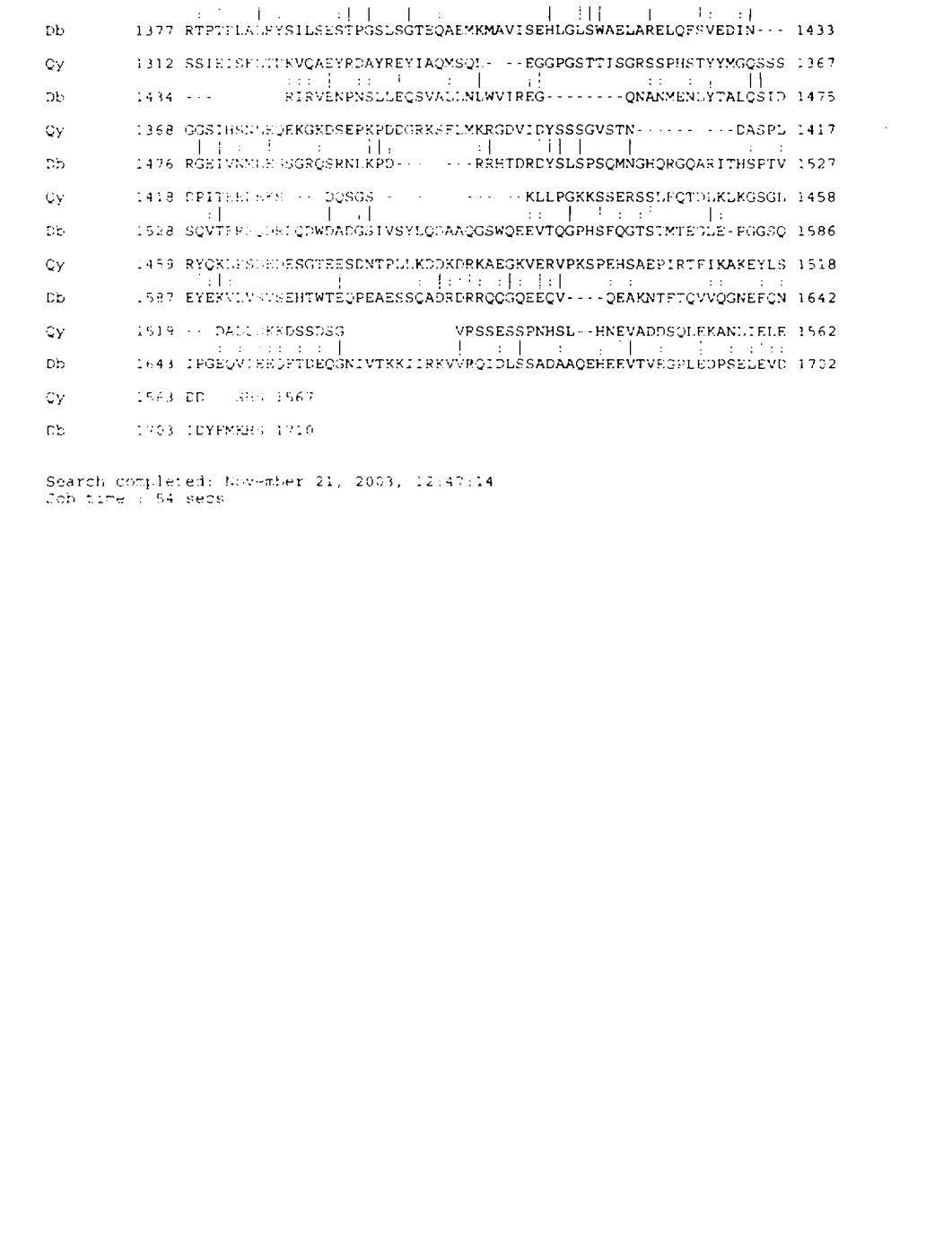


RESULT 15
Q13768
ID Q13768 PRELIMINARY; PRT; 1719 AA.
AC Q13768;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Alt ankyrin Variant 2 (.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hematopoietic;
RX MEDLINE=96158843; PubMed=8537687;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins.";
RL Nature 344:36-42(1990).
DR EMBL; X16609; CAA34611.1; -.
DR HSSP; Q00432; 1AWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000536; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1719 AA; 189011 MW; F48465D16D975CBF CRC64;

Query Match 5.7%; Score 502; DB 4; Length 1719;
Best Local Similarity 19.5%; Prob. No. 5.5e-25;
Matches 366; Conservative 229; Mismatches 653; Indels 571; Gaps 71;







Search completed: November 21, 2003, 12:47:14
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 11:43:31 ; Search time 5.5 Seconds
without alignments
941.423 Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8884
Sequence: 1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 1[illegible] residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US_PUB_MAb.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US_PUBCOMAb.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the actual database searched.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8884	100.0	1715	12	US-10-021-571-4	Sequence 4, Appli
2	8884	100.0	1715	12	US-10-021-571-2	Sequence 2, Appli
3	8830	99.4	1771	12	US-10-021-571-3	Sequence 3, Appli
4	8830	99.4	1771	12	US-10-021-571-11	Sequence 11, Appl
5	8362.5	94.1	1762	12	US-10-021-571-1	Sequence 1, Appli
6	8266	93.0	1715	12	US-10-021-571-5	Sequence 5, Appli
7	8266	93.0	1715	12	US-10-021-571-6	Sequence 6, Appli
8	5899.5	66.4	1184	12	US-10-021-571-7	Sequence 7, Appli
9	5899.5	66.4	1184	12	US-10-021-571-8	Sequence 8, Appli
10	3050.5	34.3	705	12	US-10-021-571-9	Sequence 9, Ap
11	2851.5	32.1	551	9	US-09-841-835-10	Sequence 10, Appl
12	2851.5	32.1	551	12	US-10-021-571-10	Sequence 10, Appl
13	2638	29.7	513	11	US-09-972-115A-6	Sequence 6, Appli
14	2479.5	27.8	1498	12	US-10-021-571-12	Sequence 12, Appl
15	1832.5	20.6	1398	12	US-10-021-571-13	Sequence 13, Appl
16	553	6.2	109	11	US-09-986-480-371	Sequence 371, App
17	519	5.8	1724	10	US-09-944-899-43	Sequence 43, Appl
18	474	5.3	1762	12	US-10-035-194-117	Sequence 117, App
19	437	4.9	740	9	US-09-835-788A-12	Sequence 12, Appl
20	437	4.9	740	12	US-10-175-042-12	Sequence 12, Appl
21	408.5	4.6	426	9	US-09-908-711-70	Sequence 70, Appl
22	390	4.4	1554	12	US-10-301-822-93	Sequence 93, Appl
23	385.5	4.3	765	12	US-10-128-174-3	Sequence 3, Appli
24	385.5	4.3	765	12	US-10-128-174-34	Sequence 34, Appl
25	385.5	4.3	765	12	US-10-128-174-35	Sequence 35, Appl
26	385.5	4.3	765	12	US-10-128-174-36	Sequence 36, Appl
27	385.5	4.3	765	12	US-10-128-174-37	Sequence 37, Appl
28	385.5	4.3	765	12	US-10-128-174-38	Sequence 38, Appl
29	385.5	4.3	765	12	US-10-128-174-39	Sequence 39, Appl
30	385.5	4.3	765	12	US-10-128-174-40	Sequence 40, Appl
31	385.5	4.3	765	12	US-10-128-174-41	Sequence 41, Appl
32	385.5	4.3	765	12	US-10-128-174-42	Sequence 42, Appl
33	385.5	4.3	765	12	US-10-128-174-43	Sequence 43, Appl
34	385.5	4.3	765	12	US-10-128-174-44	Sequence 44, Appl
35	384.5	4.3	1734	12	US-10-369-978-2	Sequence 2, Appli
36	379	4.3	1101	15	US-10-177-293-106	Sequence 106, App
37	376.5	4.2	285	9	US-09-835-788A-18	Sequence 18, Appl
38	376.5	4.2	285	12	US-10-175-042-18	Sequence 18, Appl
39	376.5	4.2	1327	9	US-09-841-835-2	Sequence 2, Appli
40	376.5	4.2	1327	11	US-09-972-115A-8	Sequence 8, Appli
41	376.5	4.2	1327	12	US-10-199-937-4	Sequence 4, Appli
42	376.5	4.2	1619	12	US-10-369-978-4	Sequence 4, Appli
43	370.5	4.2	949	9	US-09-841-835-10	Sequence 10, Appl
44	370	4.2	1166	11	US-09-972-115A-6	Sequence 6, Appli
45	370	4.2	1166	12	US-10-199-937-135	Sequence 135, App

ALIGNMENTS

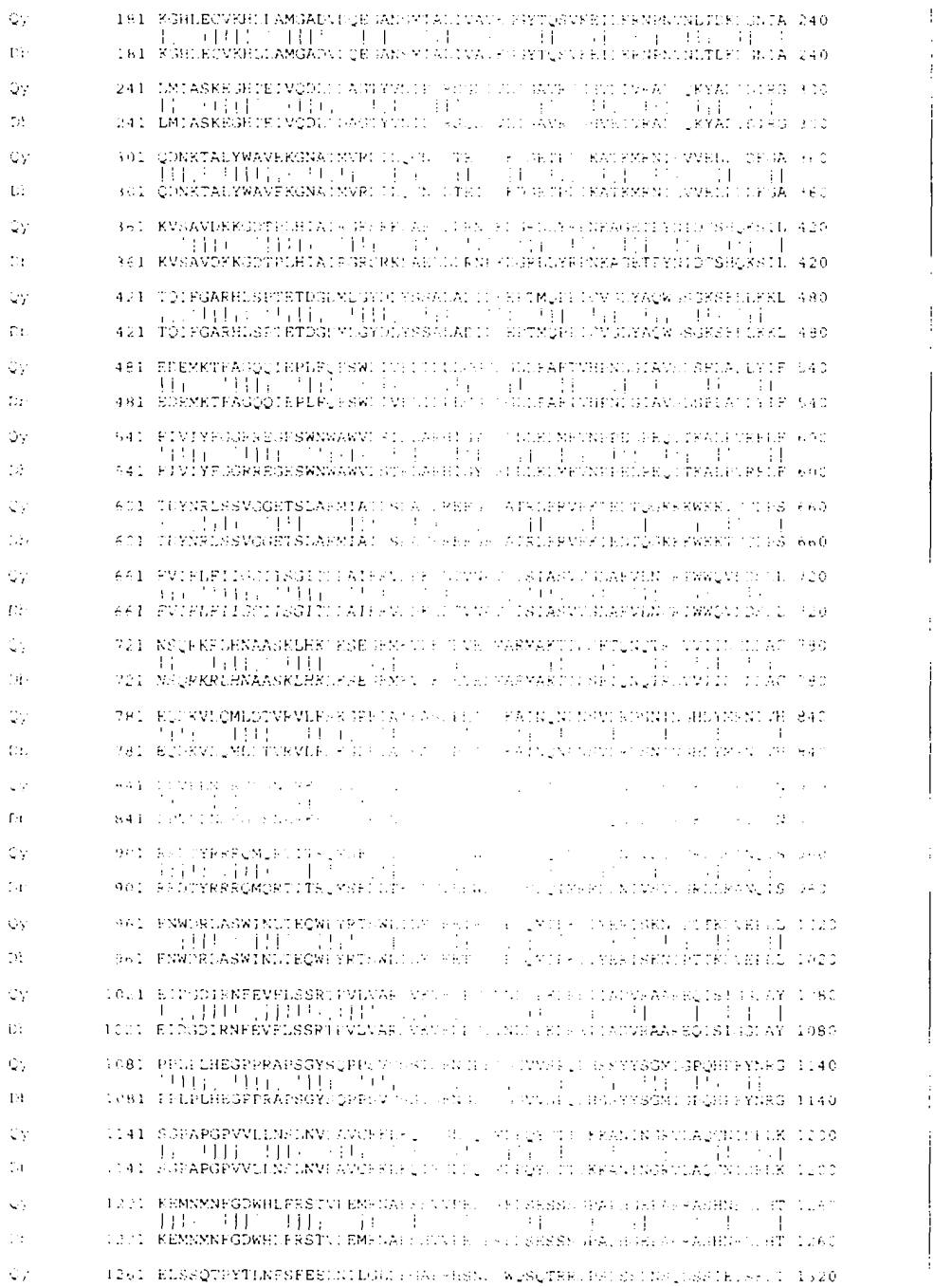
RESULT 1
US-10-021-571-4
; Sequence 4, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Dehyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL AN
; FILE REFERENCE: 8011A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,309
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-571-4

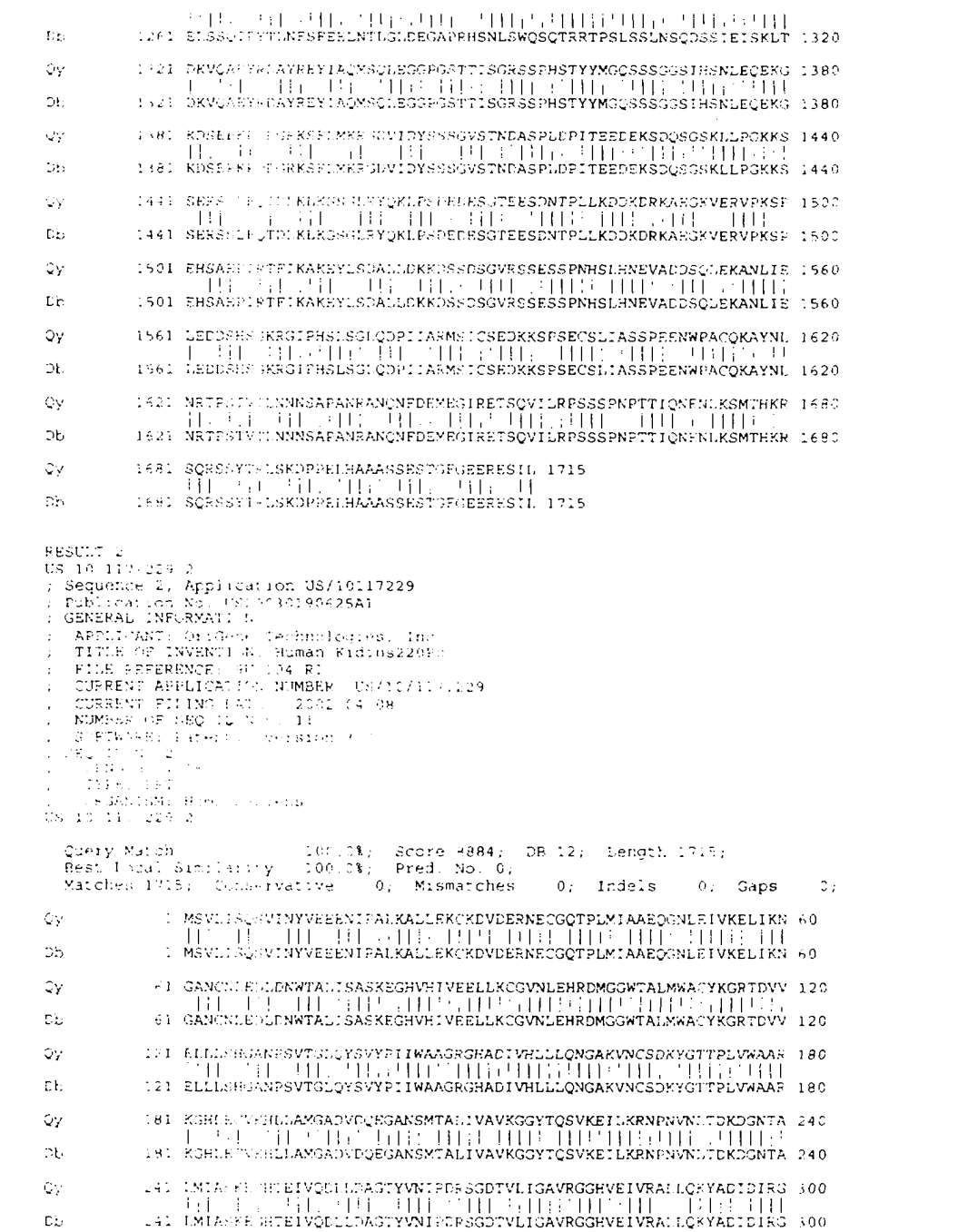
Query Match 100.0%; Score 8884; DB 12; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60
Db 1 MSVLISQSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60

Qy 61 GANCNAKMKNWTALHCASKEGHVHIVEVLLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120
Db 61 GANCNAKMKNWTALHCASKEGHVHIVEVLLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

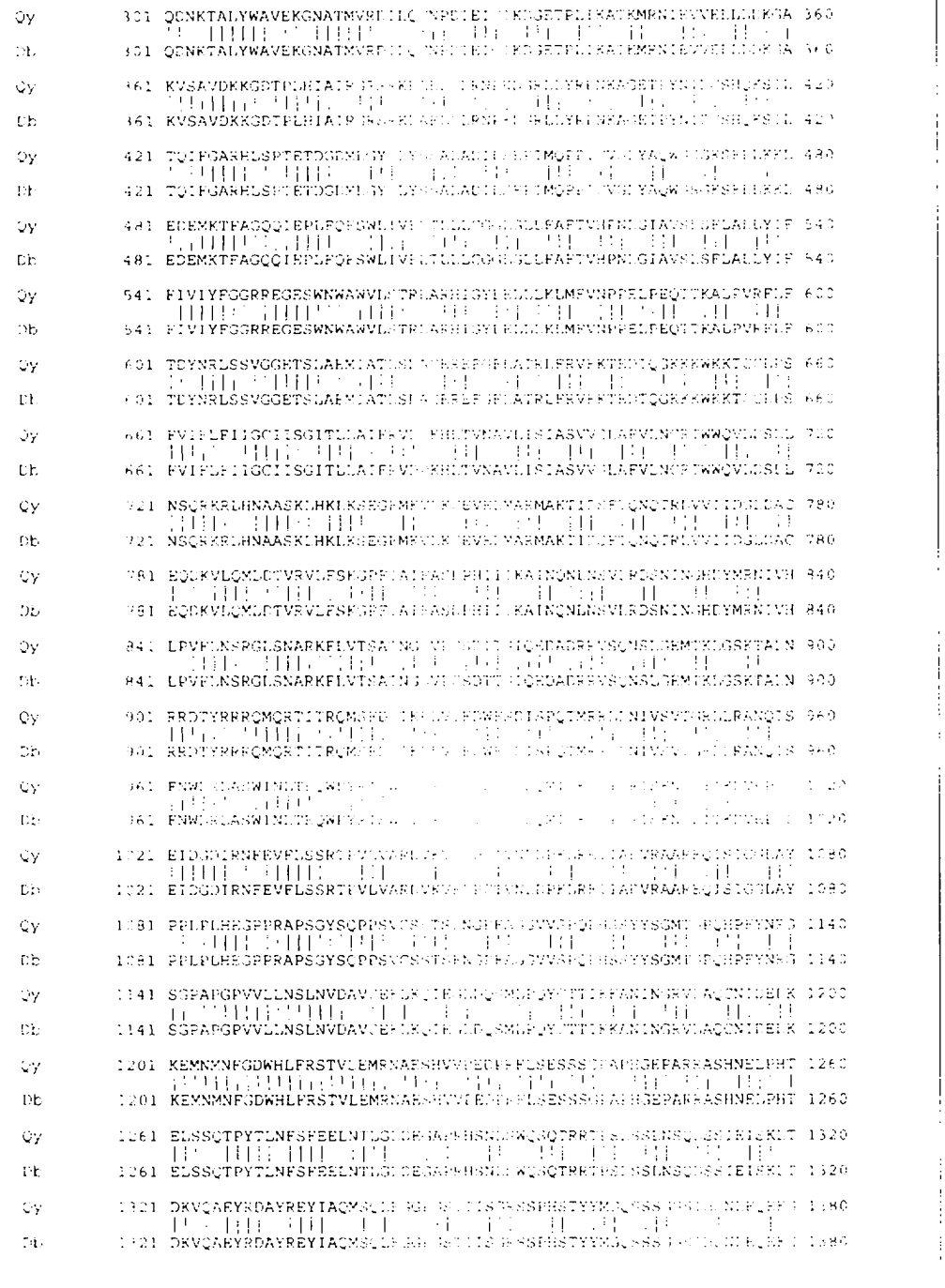
Qy 121 ELLLQ[illegible]VTHLQYSVYPFIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLL[illegible]VTHLQYSVYPFIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

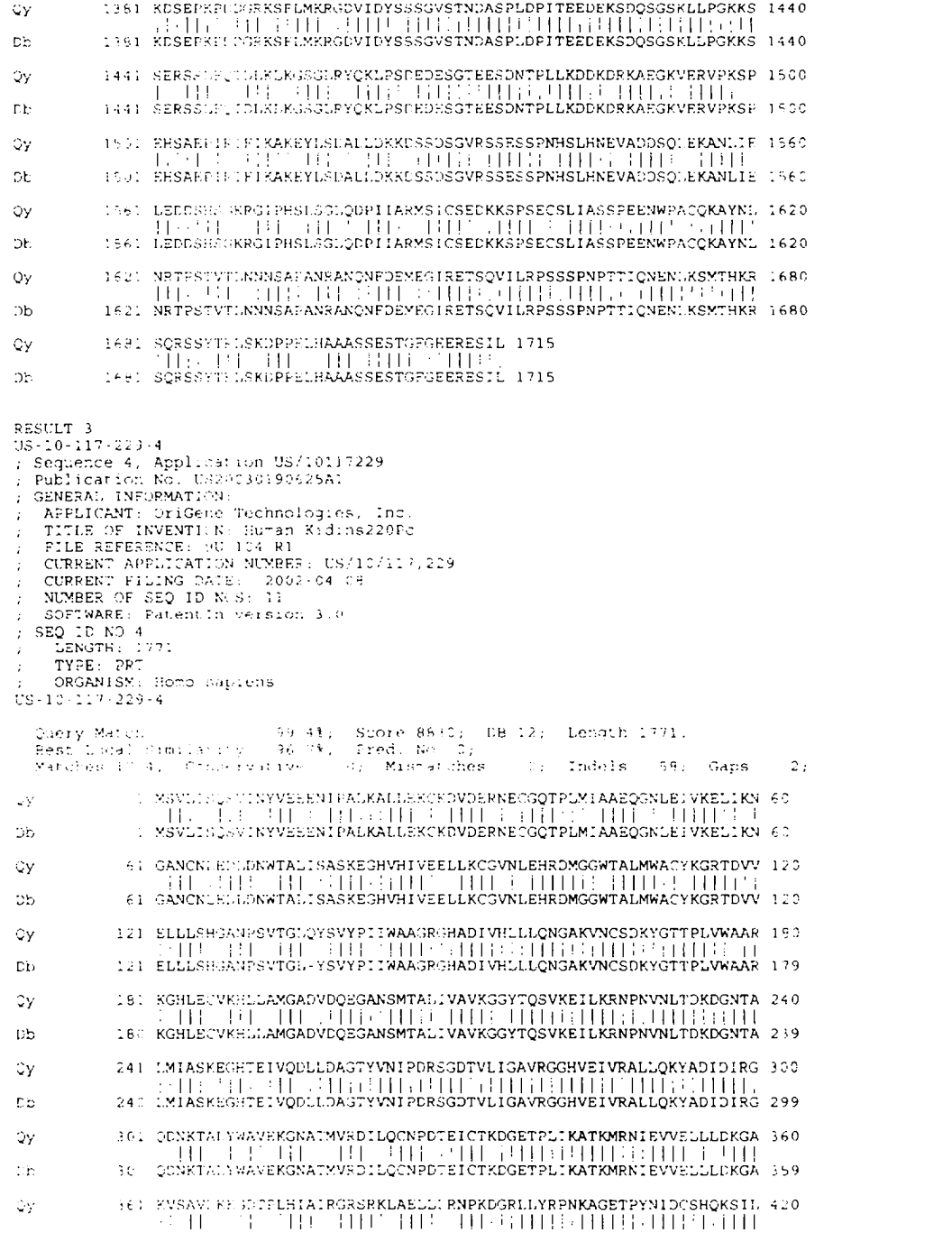




RESULT 2
US 10/10229 2
; Sequence 2, Application US/10117229
; Publication No. US20030096825A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220/ARMS
; FILE REFERENCE: 0P1004 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ IDS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US 10/10229 2

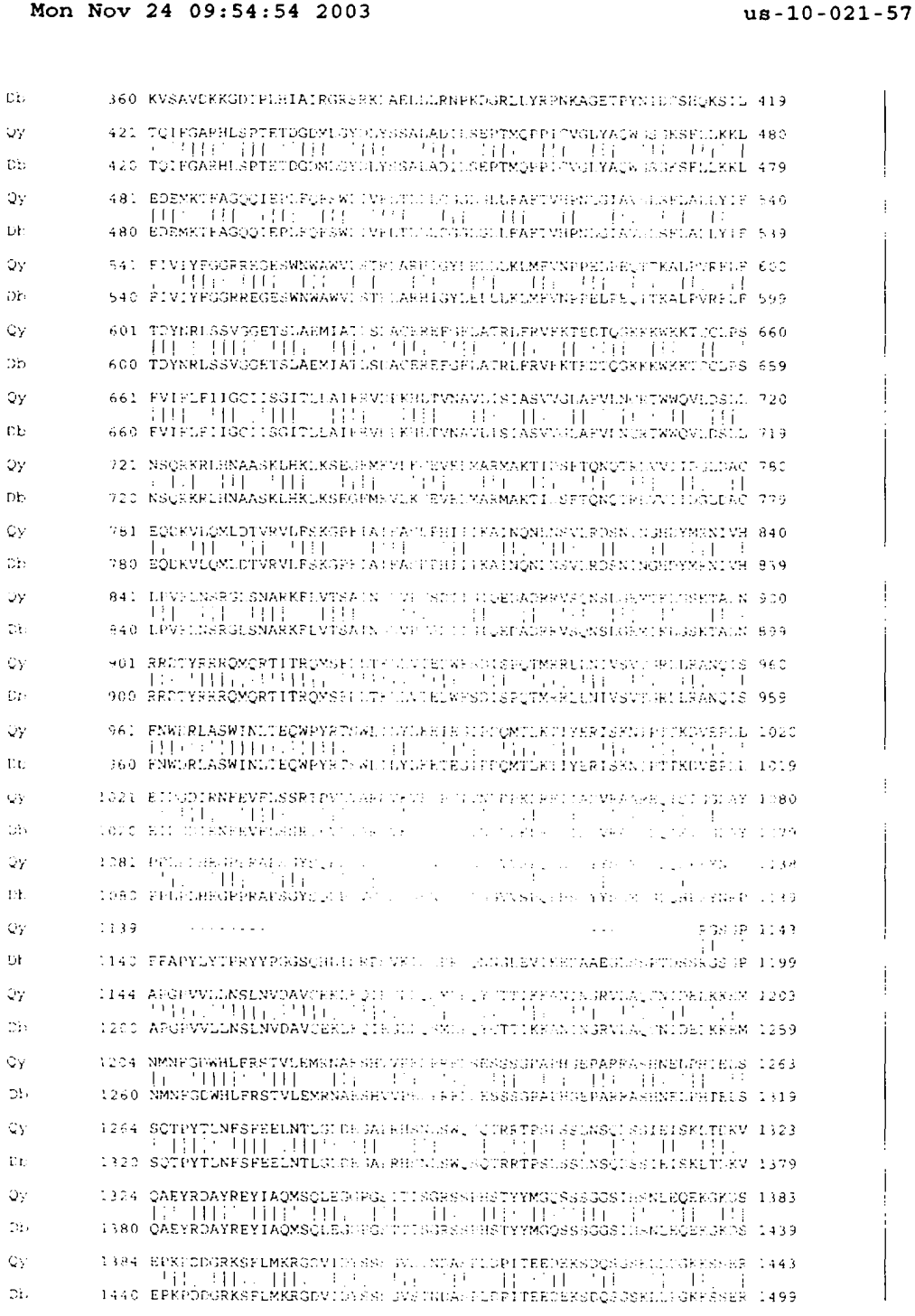
Query Match 100.0%; Score 4884; DB 12; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





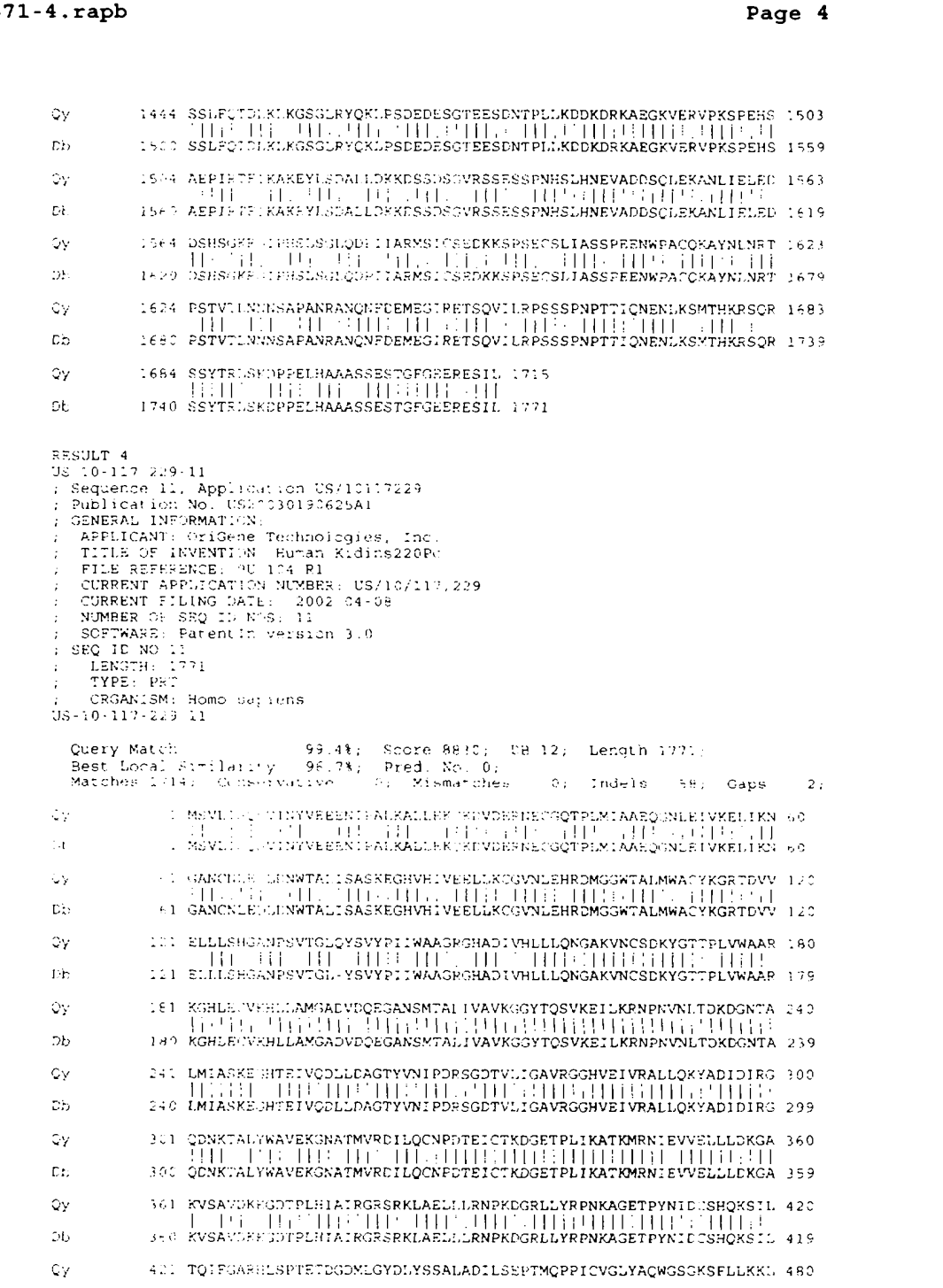
RESULT 3
US-10-117-229-4
; Sequence 4, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: OU 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-4

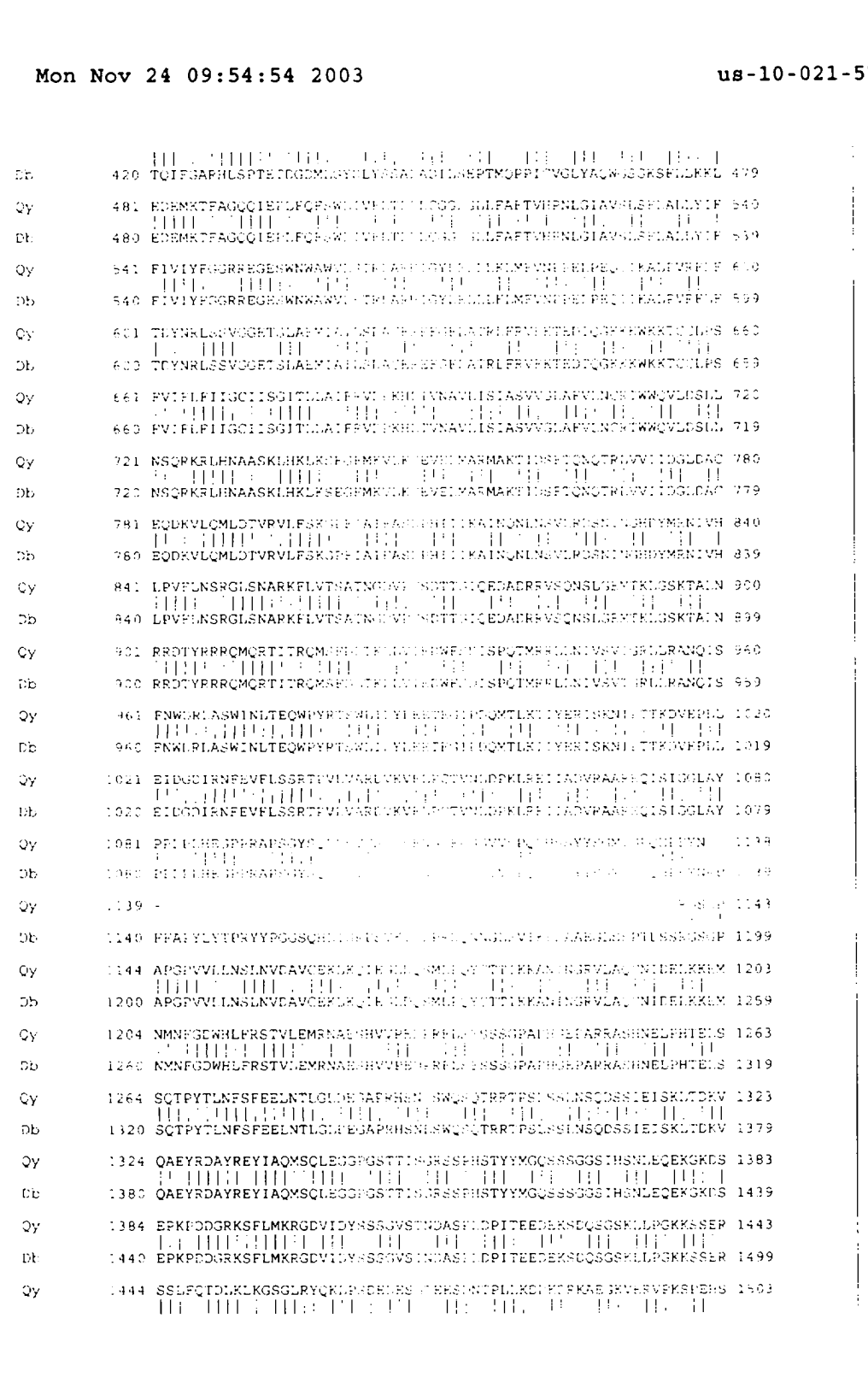
Query Match 99.4%; Score 8610; DB 12; Length 1771.
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 55; Indels 59; Gaps 2;

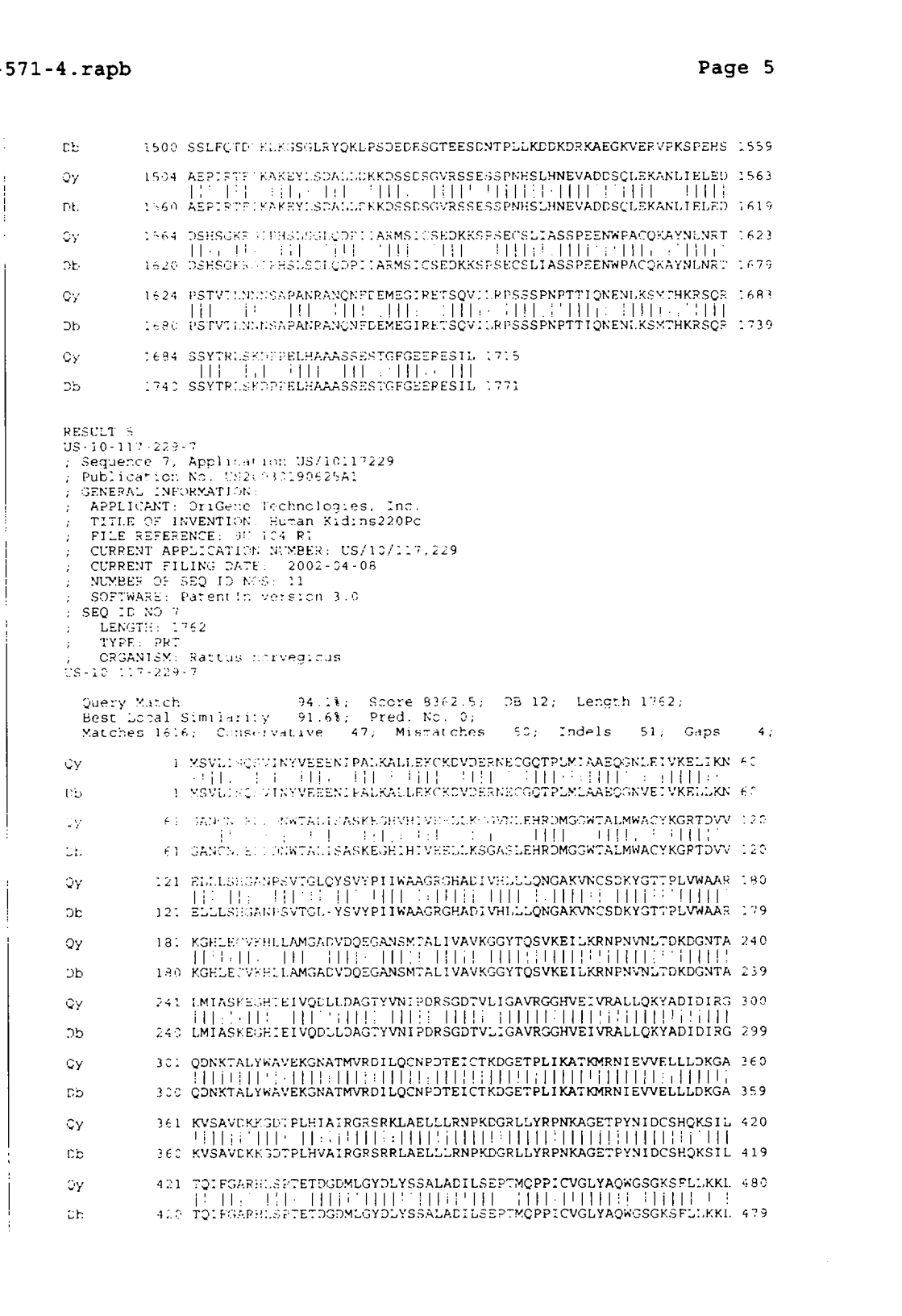


RESULT 4
US 10-117 229-11
; Sequence 11, Application US/10117229
; Publication No. US20030180625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220PC
; FILE REFERENCE: OU 104 P1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 11
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-11

Query Match 99.4%; Score 8810; DB 12; Length 1771;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

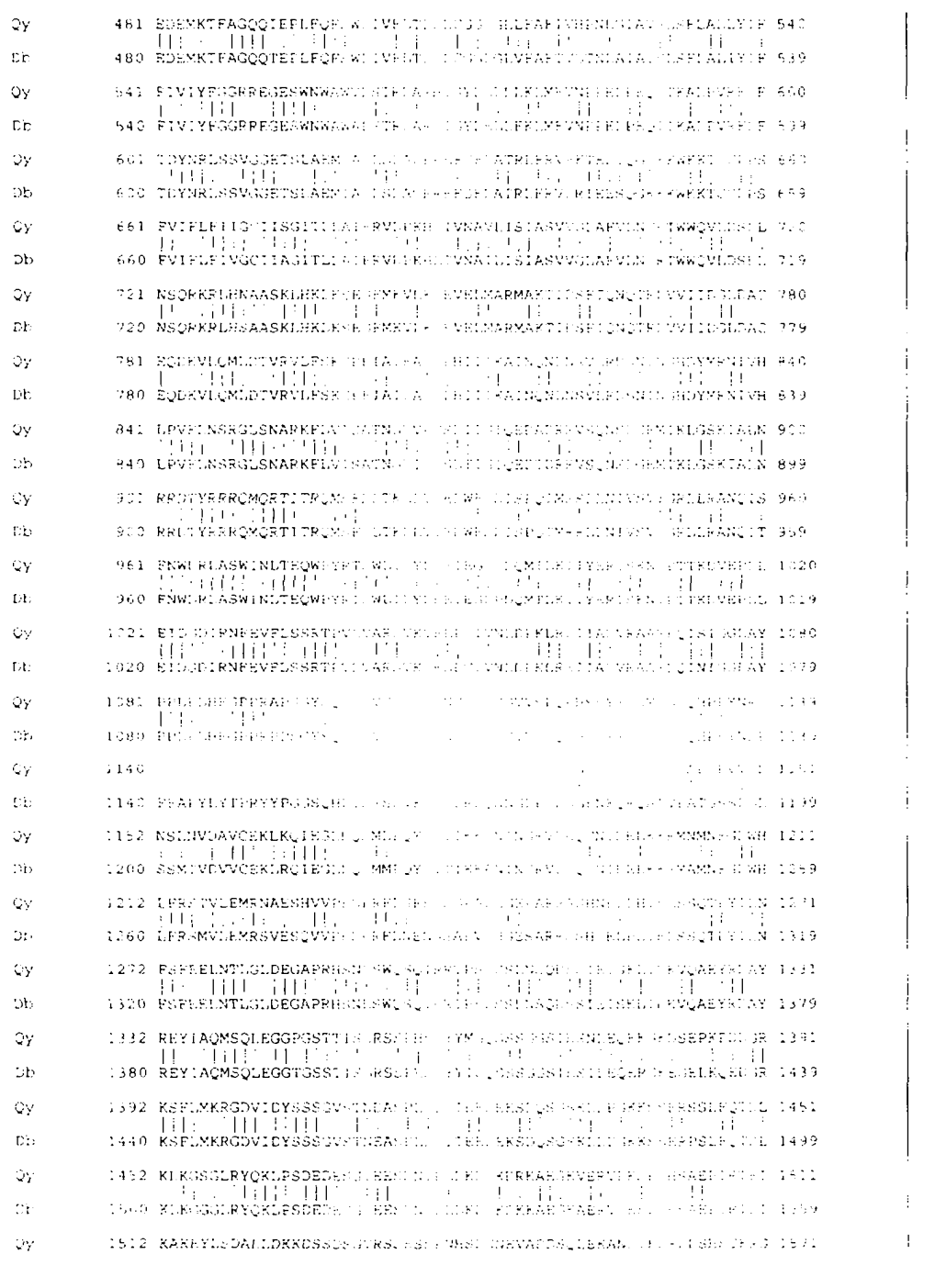


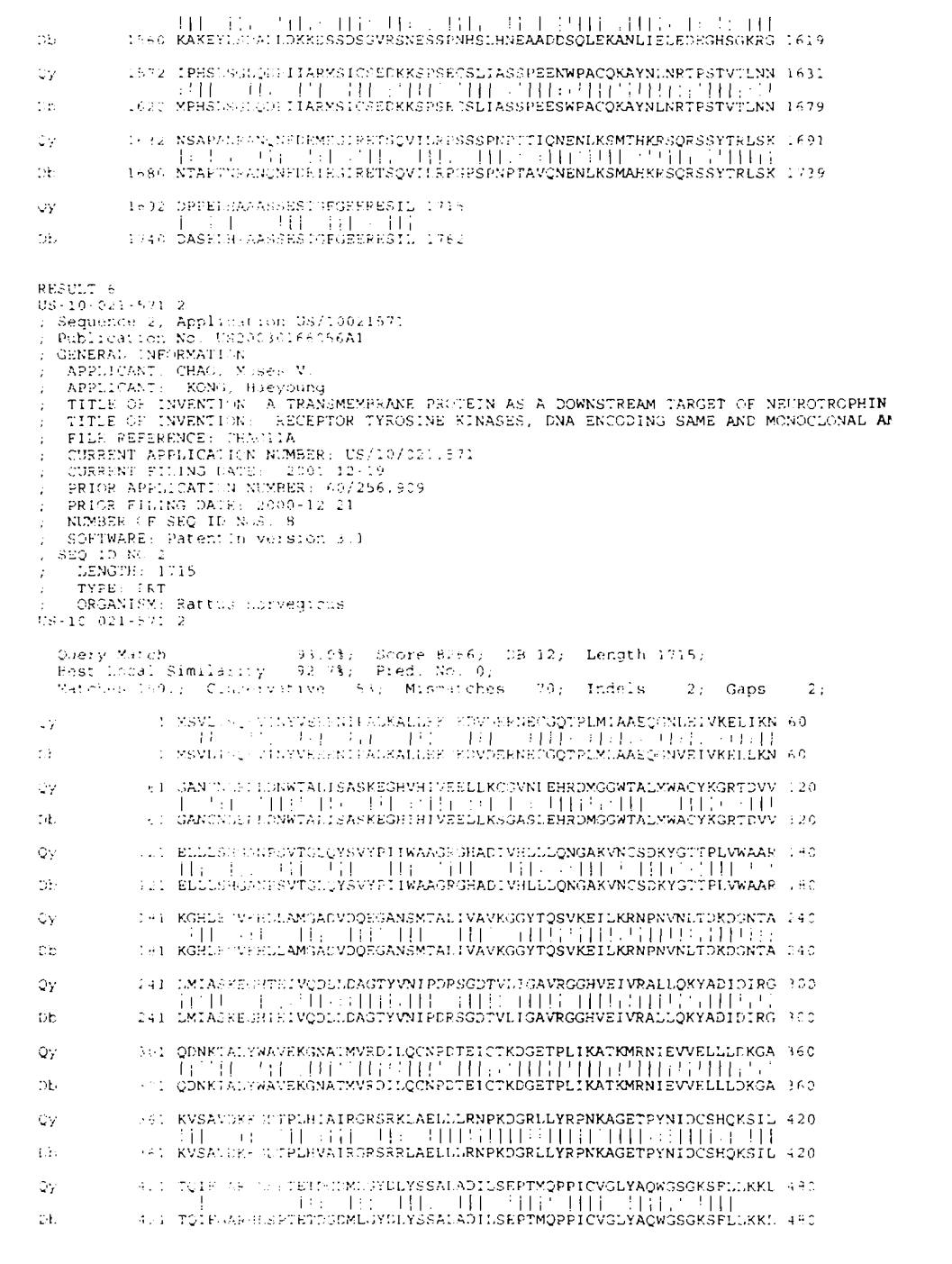




RESULT 6
US-10-117-229-7
; Sequence 7, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: OT 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 1762
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-7

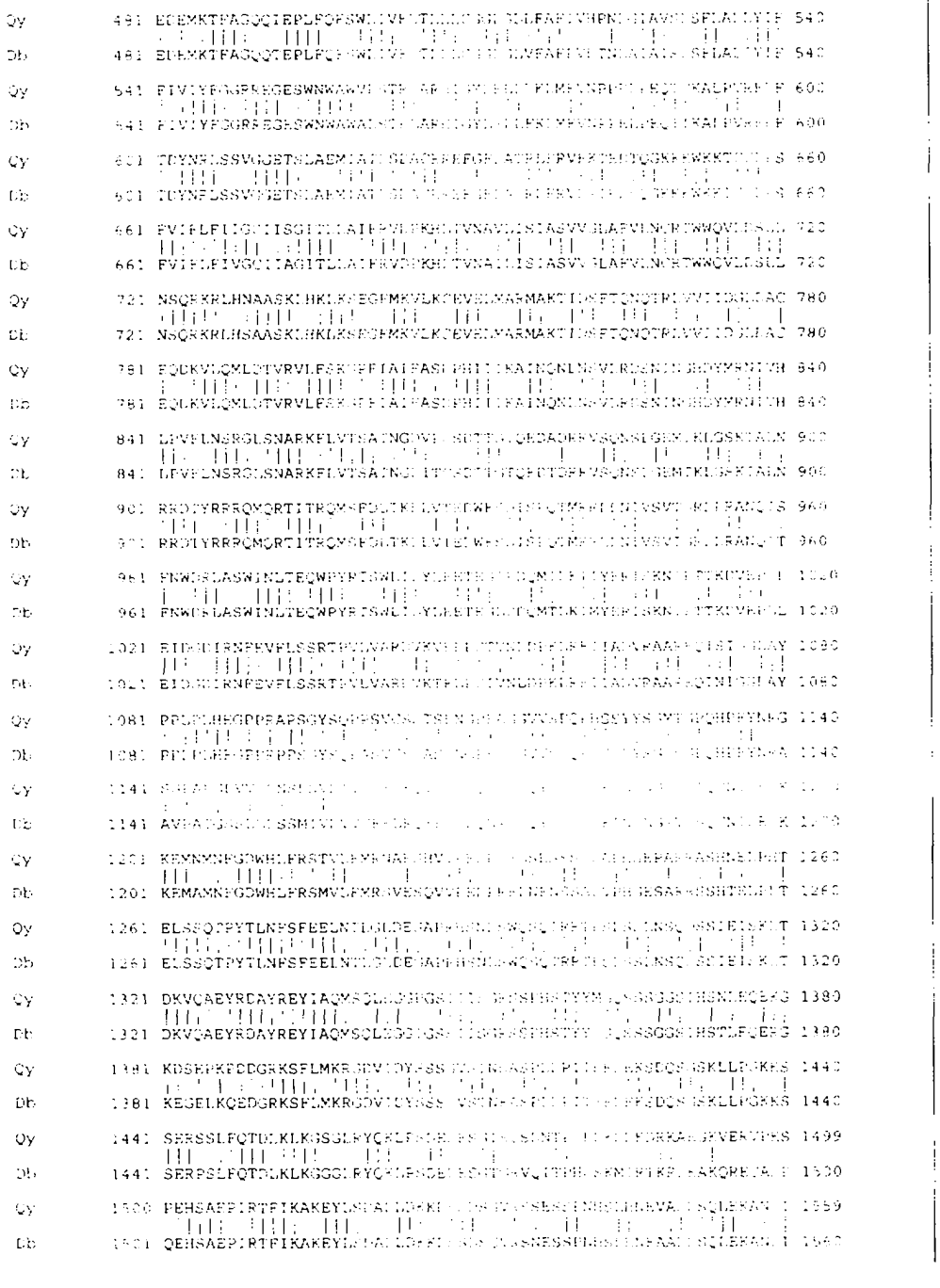
Query Match 94.1%; Score 8362.5; DB 12; Length 1762;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;





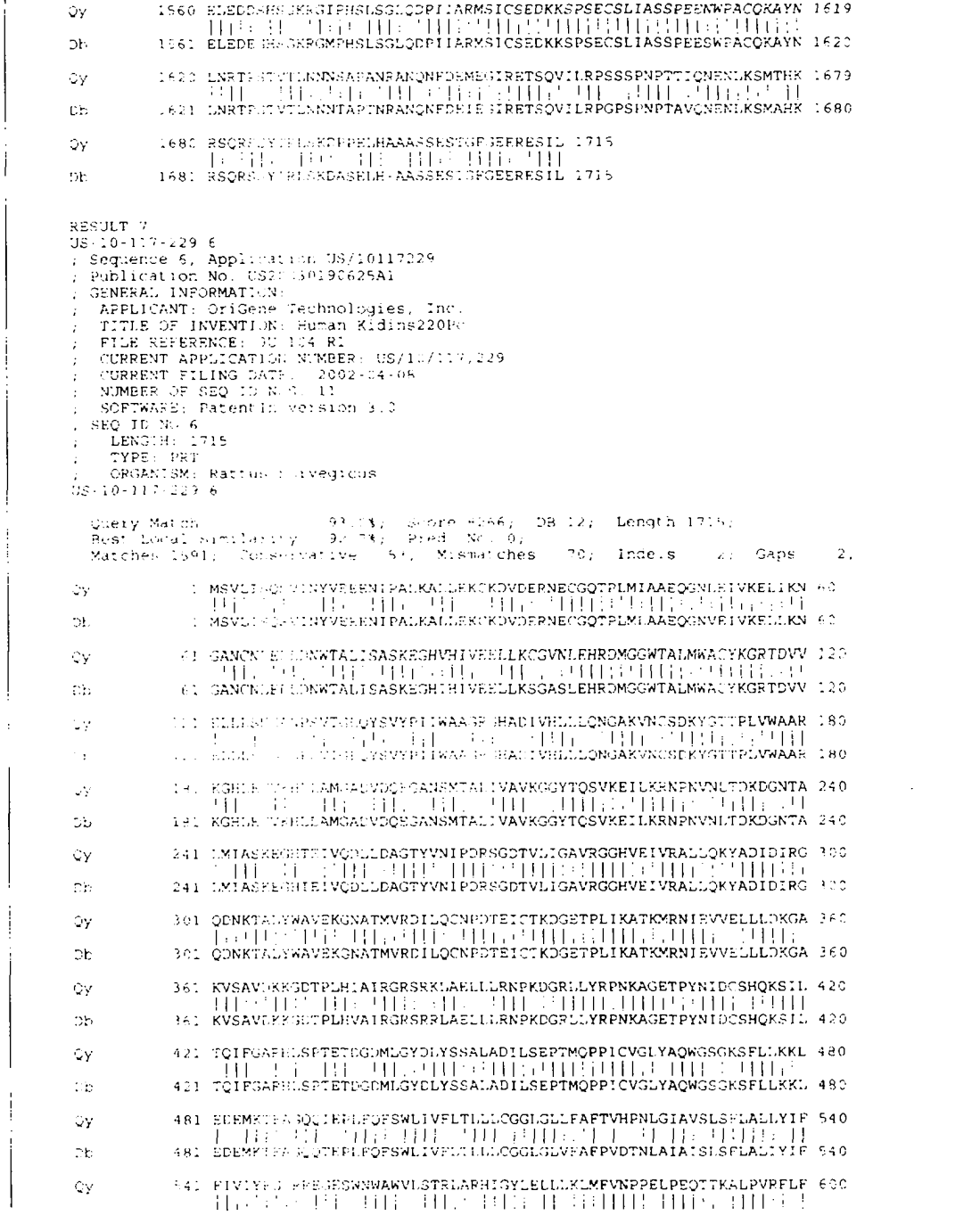
RESULT 6
US-10-021-571 2
; Sequence 2, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Hueyjoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL AN
; FILE REFERENCE: CHAO/1A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-021-571 2

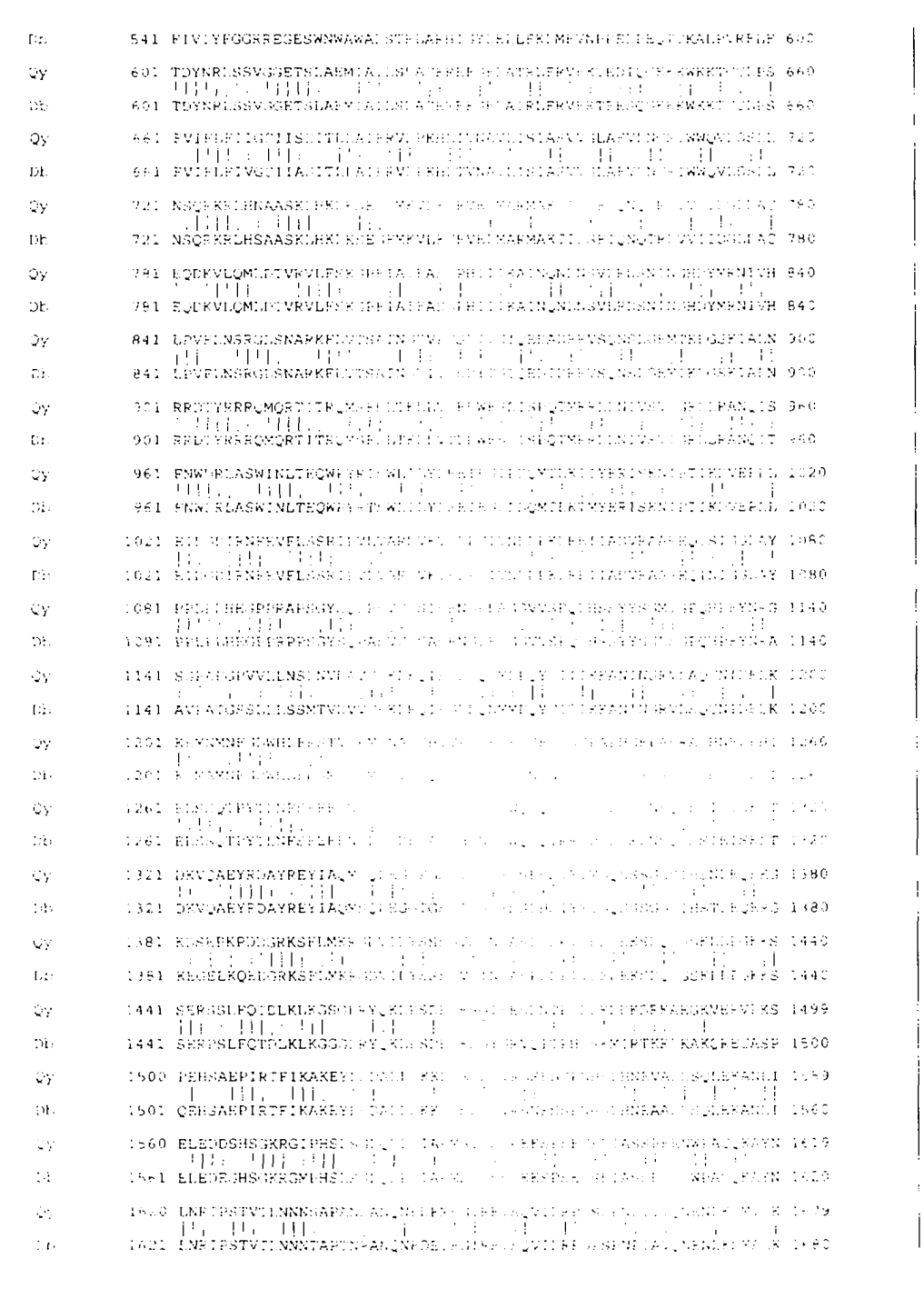
Query Match 93.6%; Score 856; DB 12; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1090; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

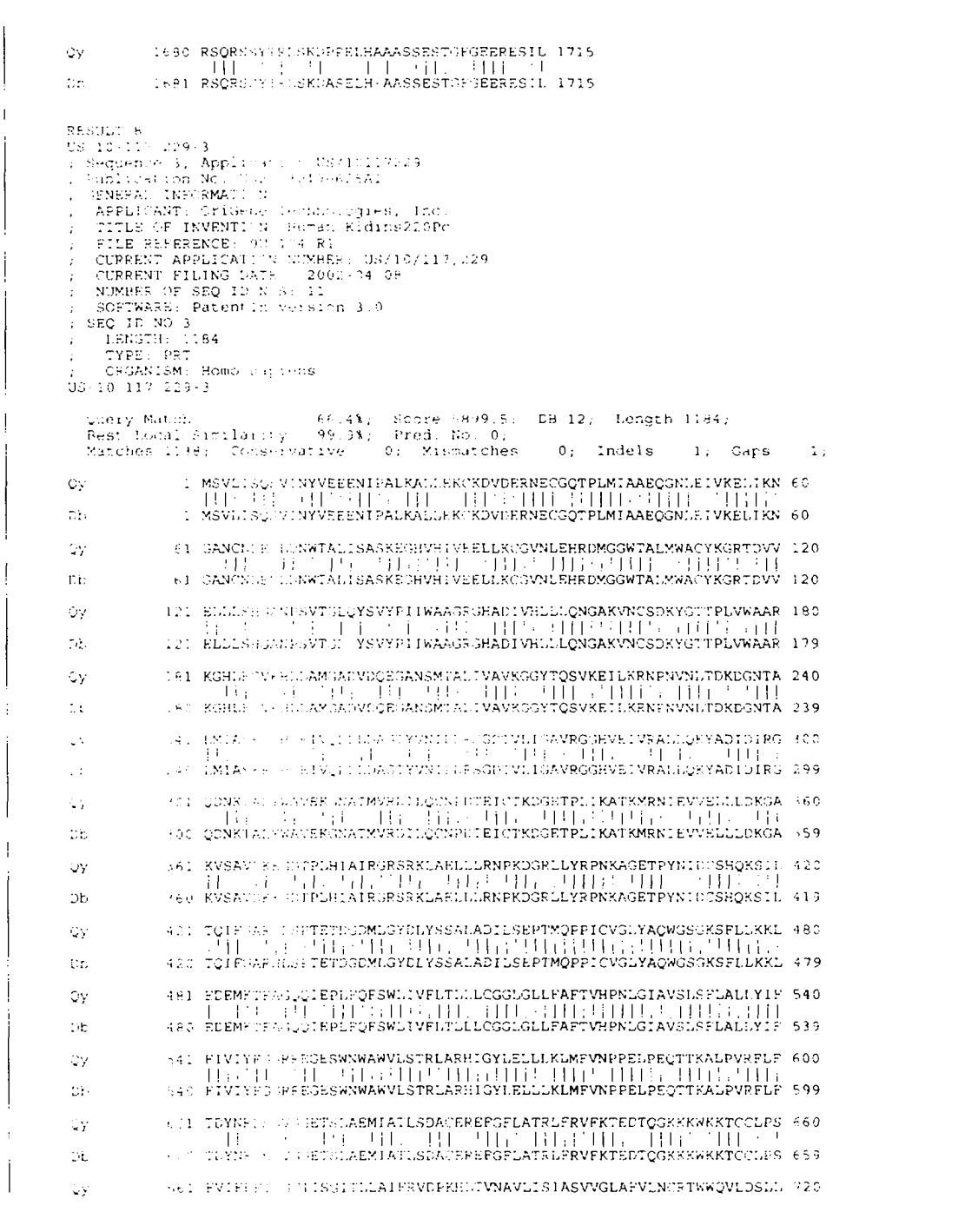


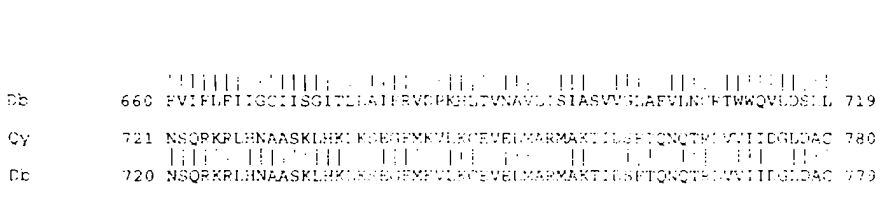
RESULT 7
US-10-117-229 6
; Sequence 6, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220PC
; FILE REFERENCE: OG 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229 6

Query Match 93.1%; Score 4366; DB 12; Length 1715;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1591; Conservative 54; Mismatches 70; Indels 21; Gaps 2;



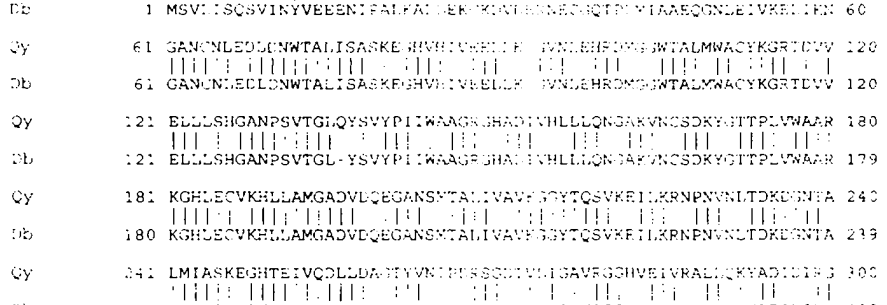


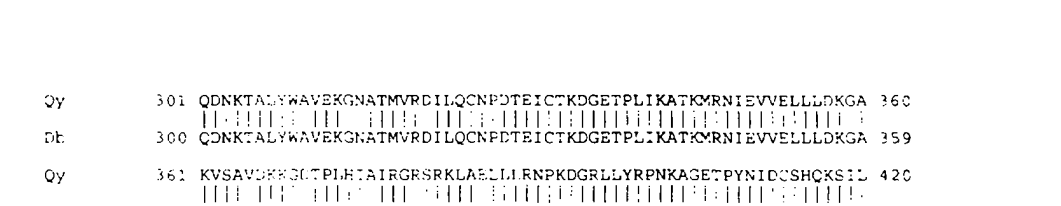




RESULT 9
US-10-117-229-9
; Sequence 9, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc.
; TITLE OF INVENTION: Human Kidney [illegible]
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-9

Query Match 66.4%; Score [illegible]; DB 12; Length 1164;
Best Local Similarity 99.9%; Pred. No. [illegible];
Matches 1138; Conservative 0; Mismatches 1; Indels 1; Gaps 1;





RESULT 10
US-10-106-698-6378
; Sequence 6378, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypepti
; FILE REFERENCE: PA006P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280

```
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6376
;    LENGTH: 705
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: MISC FEATURE
;    LOCATION: (244)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L amino acids
;    NAME/KEY: MISC FEATURE
;    LOCATION: (337)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L amino acids
US-10-106-698-6376

  Query Match             34.3%;  Score 1056.5;  DB 1;  Length 705;
  Best Local Similarity   98.5%;  Pred. No. 1.4e-204;
  Matches  587;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy         62 ANCNLEQDLGNWTALISALFPLIVVLFLLGTFGNLLVLMAVYKFRRLRSMTDVYL 121
              ::  |||  ||||  |||:  |:|:  :|:|  |||  |::  |:|  |':  ||.
Db         26 SNCNLEDDLGNWTALISALFPVIVVLFLLGTFGNLLVIMAVYKFRRLRSMTDVYL 85

Qy        122 LNLSIADFLFLVTLPFWAHSAANQWVFGNAMCKLFTGLYHIGYFGGIFFIILLTID 181
              |:| :|||: |||  : :|||  : |  :|  :|  :|  ||  :||  :|:|
Db         86 LNLSIADFLFLVTLPFWAHSAANQWVFGNAMCKLFTGLYHIGYFGGIFFIILLTID 145

Qy        182 GHLECVKHLLAMGADVDQEGANSMTALIVAVKQGYTQSVKALLDAGTYVNIPDRSGD 241
              ||||:|||:|||  |:||  |:||  |:  :|:  |:|:|:|||:|||  ||  ||
Db        146 GHLECVKHLLAMGADVDQEGANSMTALIVAVKQGYTQSVKALLDAGTYVNIPDRSGD 205

Qy        242 MIASKEGHTEIVQLLLKAGVNIDEKNKDFDSPLHEAAKYGYIEIVKALLEYGADIEIQ 301
              | :||  |||  |||:  ||:  :|:|  ||  :|  :|  :|  ||  |:|  |
Db        206 MIASKEGHTEIVQLLLKAGVNIDEKNKDFDSPLHEAAKYGYIEIVKALLEYGADIEIQ 265

Qy        302 DNKICALYWAVEKGNATMVRILQCNPDTEICTKDGETPLDIAIRYRRSRKLAELLLRN 361
              |||:|||  |||  ||:  |:|  :|:  :|:  ||  |||  ||:|
Db        266 DNKICALYWAVEKGNATMVRILQCNPDTEICTKDGETPLDIAIRYRRSRKLAELLLRN 325

Qy        362 VNAVDKKGDTPLHIAIRNKRPLAVLLIRDGF  392
              ||||:  :| ||  |:|  :|:  :|  |:|
Db        326 VNAVDKKGDTPLHIAIRNKRPLAVLLIRDGF  356

Qy        393 QIRGAHELVRNDCSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQ 452
              |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|
Db        357 QIRGAHELVRNDCSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQ 416

Qy        453 PPIGVYAQWPSGKSFLLKKLEDEMKTPAQQQIEPLFQFSWLIVPLTLLLCGGLGLLFA 512
              |||  |:  |||  |||  |||  :|||  |||  |:|  |:|  |:|  |:|  |
Db        417 PPIGVYAQWPSGKSFLLKKLEDEMKTPAQQQIEPLFQFSWLIVPLTLLLCGGLGLLFA 476

Qy        513 FTVMSEYVKLNKQLEYPTRVTFNMLESSGYRSGGSSGWVPSLRLCVQLQERHIGYLD 572
              |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |
Db        477 FSVMSEYVKLNKQLEYPTRVTFNMLESSGYRSGGSSGWVPSLRLCVQLQERHIGYLD 536

Qy        573 LLTVFLMCSLQIVIMAGFVGLFIYGCFWYFVRSKVLAVELGDKVERSGSFLPSFHSL 632
              |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |
Db        537 LLTVFLMCSLQIVIMAGFVGLFIYGCFWYFVRSKVLAVELGDKVERSGSFLPSFHSL 596

Qy        633 DYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLRRVF 672
              |:  |||:  |||:  |||  :|:|  :|  ||:  |:  |:|
Db        597 DYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLRRVF 636

RESULT 11
US-09-835-788A-17
; Sequence 17, Application US/09/835,788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides,
;  TITLE OF INVENTION:  Antibodies
;  FILE REFERENCE: PT019P1
;  CURRENT APPLICATION NUMBER: US/09/835,788A
;  CURRENT FILING DATE:  2001-04-17
;  PRIOR APPLICATION NUMBER: PCT/US00/28666
;  PRIOR FILING DATE: 2000-10-17
;  PRIOR APPLICATION NUMBER: 60/159,585
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/167,246
;  PRIOR FILING DATE: 1999-11-24
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;    LENGTH: 551
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-835-788A-17

  Query Match             32.1%;  Score 988.5;  DB 9;  Length 551;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-192;
  Matches  550;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        102 MGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNG 161
              |||  |:|  ||| .'|||  ||| : ||||  ||||||||||||.||||.':|||||:
Db          1 MGGWTAIMWACYKGRTDVVELLLSHGANPSVTGL-YSVYPIIWAAGRGHADIVHLLLQNG 59

Qy        162 AKVNCGLRYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKQGYTQSVK 221
              :||||  ||.'|||  |||.'|||:.||||||||||||||||:|||| ;'||||
Db         60 AKVNCGLRYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKQGYTQSVK 119

Qy        222 EILKKNPDVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 281
              |||:||, |||  |:|, |||: ||||:|||||:|||| |||||:|||| | |||  .||
Db        120 EILKKNPDVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGG 179

Qy        282 HVEIVKALLEYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLD 341
              |||  .:|  ||. |||  |||.: ||||  ||||:|||||.|||||  |||||:
Db        180 HVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLD 239

Qy        342 KATKMPNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRP 401
              ' |||  :'.  |||  :|||, ||||, ||||.:||||,,'|||,' ||||:|:|, '|||
Db        240 KATKMPNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRP 299

Qy        402 NKAGETFYNIDCSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 461
              |.:|||  ||,' ||,  |||  :||  :|||,'||||||:!||||  ||||, '|||| :|
Db        300 NKAGETFYNIDCSHQKSILTQIPGARHLSPTETDGDMLGYDLYSSALADILSEPTMQPPI 359

Qy        462 GVGIYAQWPSGKSFLLKKLEDEMKTPAQQQIEPLFQFSWLIVPLTLLLCGGLGLLFAFTV 521
              |||  |'  |||  |||  ||,  '|||:. |||| :||||:| |||| ' ||||| |
Db        360 GVGLYAQWPSGKSFLLKKLEDEMKTPAQQQIEPLFQFSWLIVPLTLLLCGGLGLLFAFTV 419

Qy        522 HFNL:'.  :'  :'PLALLYIFFIVYFDPPREGESWNWAWVLSTPLARHIGYLFLLLKLMFV 581
              ||      |':  |:|  ||.  |||  :|||  :|||.  |||| :'|||
Db        420 HFNLVFA. ISFLALLYIFFIVYFDPPREGESWNWAWVLSTPLARHIGYLFLLLKLMFV 479

Qy        582 NPPELFE.:TFKALPVRFLFTDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLRRVF 641
              |:||||  ':|| :|||  ' |||  ''|||  :|||||:|||||. ||||.  ||||  ||
Db        480 NPPELFE.QTFKALPVRFLFTDYNRLSSVGGETSLAEMIATLSDACEREFGFLATRLRRVF 539

Qy        642 KTEDT:PFFKWK 653
              |||''||,  !
Db        540 KTEDI:PFFKKK 551


RESULT 12
US-10-175-042-17
; Sequence 17, Application US/10175042
; Publication No. US20030181710A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides,
;  TITLE OF INVENTION:  Polypeptides, and
;  TITLE OF INVENTION:  Antibodies
;  FILE REFERENCE: PT019P1
;  CURRENT APPLICATION NUMBER: US/10/175,042
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: 09/835,788
;  PRIOR FILING DATE:  2001-04-17
;  PRIOR APPLICATION NUMBER: PCT/US00/28666
```

; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/[illegible],56
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/[illegible],24
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US 10/175-042-17

Query Match 32.1%; Score 1060.5; DB 12; Length 551;
Best Local Similarity 99.6%; Pred. No. 1.1e-192;
Matches 550; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 102 MGGWTALMWACYKGRTDVVELLLRHGAEVSVI[illegible] 161
Db 1 MGGWTALMWACYKGRTDVVELLLRHGAEVSVD[illegible] 59

Qy 162 AKVNGSDKYGSTPLVWAAAFGHLEVVKFLLAY[illegible] 221
Db 60 AKVNGSDKYGSTPLVWAAAFGHLEVVKFLLAY[illegible] 119

Qy 222 EILKPNPNVNLTDKDGNTALYIASRE[illegible] 281
Db 120 EILKPNPNVNLTDKDGNTALYIASRE[illegible] 179

Qy 282 HVEIVKALLQKYADIDIR[illegible] 441
Db 180 HVEIVKALLQKYADIDIR[illegible] 239

Qy 342 KATPMRNIPVVELLLDKGAKVNAVMHGTPLIAIR[illegible] 401
Db 240 KATPMRNIPVVELLLDKGAKVNAVTKHGTPLIAIR[illegible] 299

Qy 402 NKAGETPYNIDGSHQKSIIQLPPAMPLPTS[illegible] 461
Db 300 NKAGETPYNIDGSHQKSIIQLPPAMPLPTS[illegible] 359

Qy 462 CVGETALMQLIKSDPIIEK[illegible] 521
Db 360 CVGETALMQLIKSDPIIEK[illegible] 419

Qy 522 HPNLGIAVSISFLALDYIPFGYFYGLPSLGWAVS[illegible] 581
Db 420 HPNLGIAVSISFLALDYIPFGYFYGLPSLGWAVTS[illegible] 479

Qy 582 NPPELPEQTTKALPVRFLPIFNRDGSYGGEIAEMIAI[illegible] 641
Db 480 NPPELPEQTTKALPVRFLPIFNRDGSYGGEIAEMIAI[illegible] 539

Qy 642 KTEDTQGKKKWK 653
Db 540 KTEDTQGKKKKK 551

RESULT 13
US 10-149-819-9
; Sequence 9, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandr.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0740 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-20; 1999-12-16
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: US20030044913A1 2215706CD1
US 10-149-819-9

Query Match 29.7%; Score 2638; DB 15; Length 513;
Best Local Similarity 99.8%; Pred. No. 1.6e-177;
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1203 MNMNPGDWHLPRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 1262
Db 1 MNMNPGDWHLPRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 60

Qy 1263 SSQTPYTNPSPEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDK 1322
Db 61 SSQTPYTNPSPEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNSQDSSIEISKLTDK 120

Qy 1323 VQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKGKD 1382
Db 121 VQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGGSIHSNLEQEKGKD 180

Qy 1383 SEPKPDGRKSFLMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE 1442
Db 181 SEPKPDGRKSFLMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE 240

Qy 1443 RSSLFQIILKLKGSGLRYQKLPSDEDESGTEESDNTPLLKDDKDRKAEGKVERVPKSPEH 1502
Db 241 RSSLFQIILKLKGSGLRYQKLPSDEDESGTEESDNTPLLKDDKDRKAEGKVERVPKSPEH 300

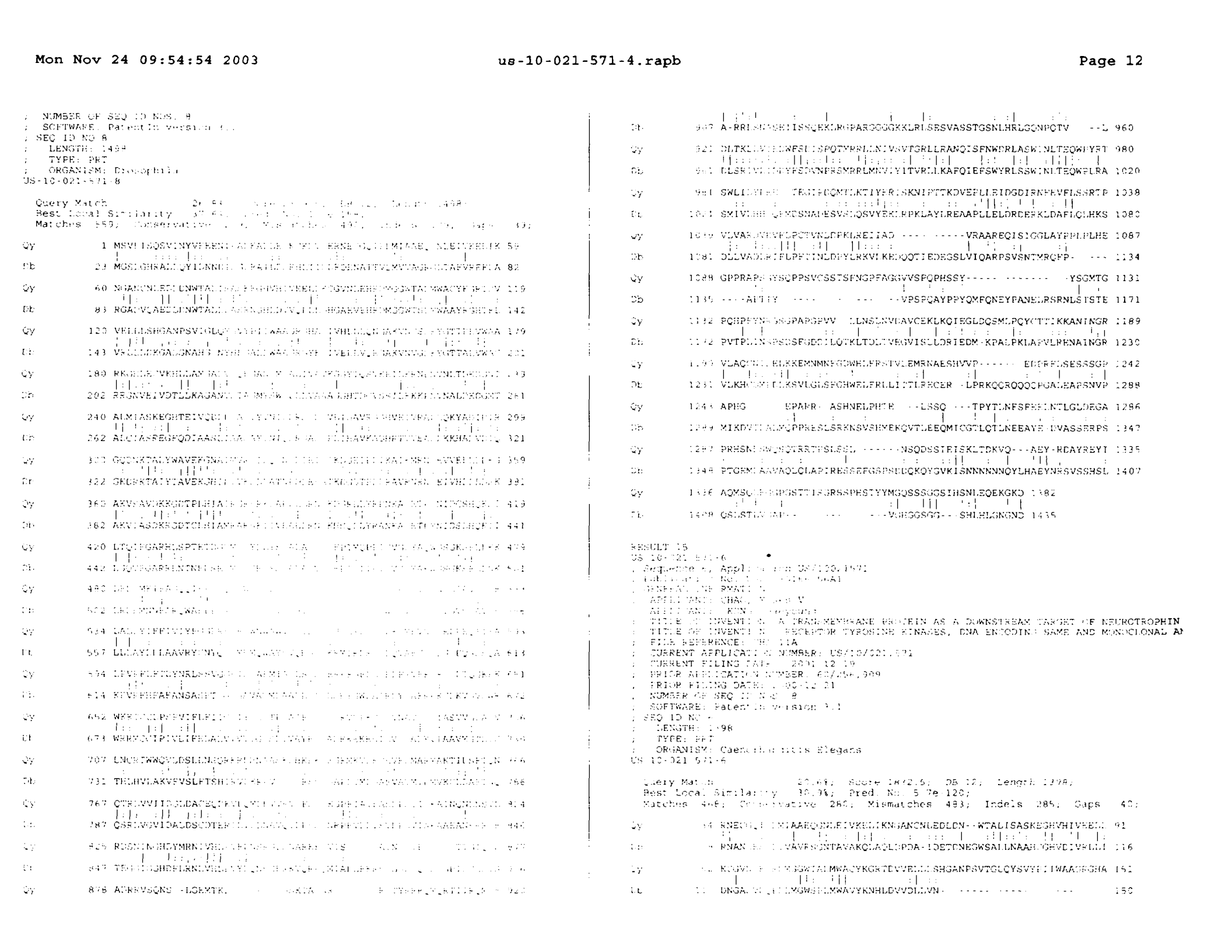
Qy 1503 [illegible] 1562
Db 301 [illegible] 360

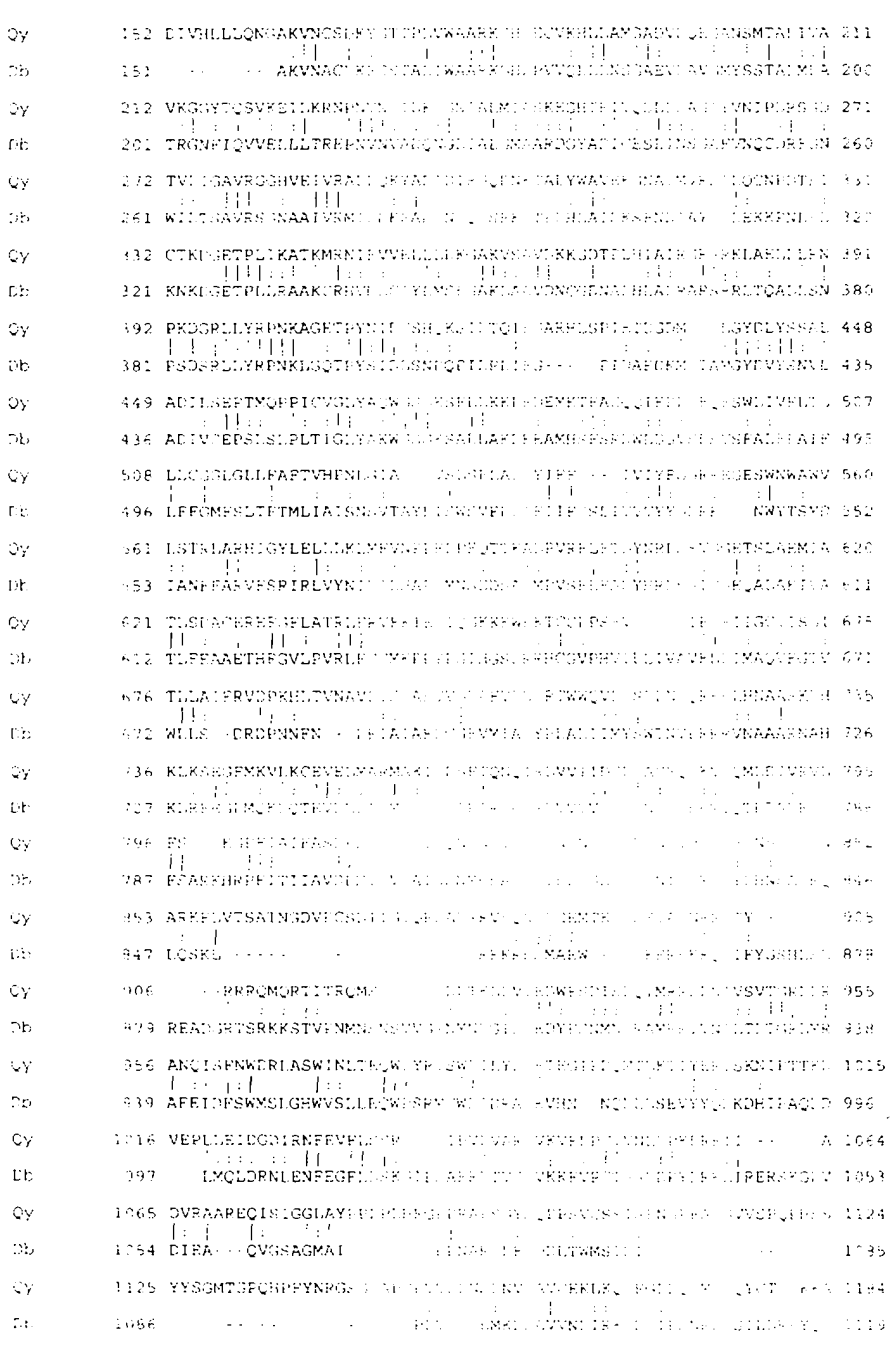
Qy 1563 DDSHSQFSIFHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNR 1622
Db 361 DDSHSQFSIFHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNR 420

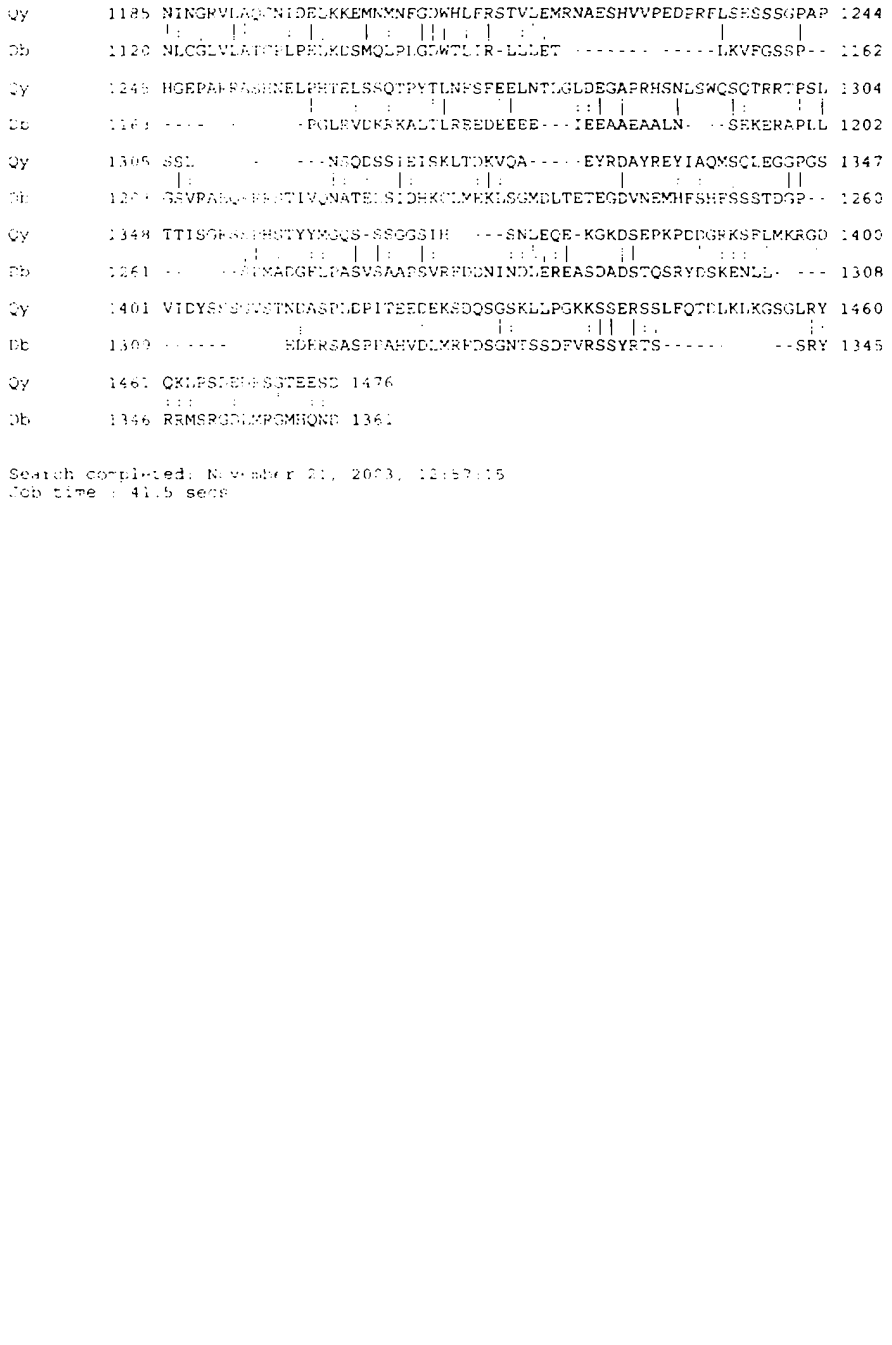
Qy 1623 TPSTVTLQNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRSQ 1682
Db 421 TPSTVTLQNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNENLKSMTHKRSQ 480

Qy 1683 RSSYTSLGKDPPELHAAASSESTGFGEERESIL 1715
Db 481 RSSYTSLGKDPPELHAAASSESTGFGEERESIL 513

RESULT 14
US-10-021-571-8
; Sequence 8, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Haeyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL AN
; FILE REFERENCE: "HA 11A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21







Search completed: November 21, 2003, 12:57:15
Job time : 41.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 12:47:53 ; Search time 1350.33 Seconds
(without alignments)
481.111 Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8884
Sequence: 1 MSVLISQSVINYVPENIPA ... LHAAAAGPGIETPDTPKQSIL 1839

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 423108694 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/PCT_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/iaa/PCT_US_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	530.5	6.0	1839	4	US-09-172-977-4	Sequence 4, Appl
2	493.5	5.6	894	2	US-08-436-771-2	Sequence 2, Appli
3	480.5	5.4	1746	2	US-09-172-977-33	Sequence 33, Appl
4	480.5	5.4	1746	2	US-09-065-474-33	Sequence 33, Appl
5	480.5	5.4	1746	3	US-09-065-474-33	Sequence 33, Appl
6	480.5	5.4	1746	4	US-09-557-034-33	Sequence 33, Appl
7	438	4.9	1088	4	US-09-082-[illegible]-2	Sequence 2, Appli
8	409.5	4.6	1423	2	US-08-810-712-10	Sequence 10, Appl
9	394.5	4.4	1704	4	US-09-392-810-2	Sequence 2, Appli
10	376.5	4.2	1327	2	US-09-172-977-3	Sequence 3, Appli
11	376.5	4.2	1327	4	US-09-641-841-3	Sequence 3, Appli
12	376.5	4.2	1619	4	US-09-392-810-4	Sequence 4, Appli
13	374	4.2	348	2	US-09-031-495-28	Sequence 28, Appl
14	374	4.2	348	2	US-08-847-429-28	Sequence 28, Appl
15	374	4.2	348	3	US-09-065-474-28	Sequence 28, Appl
16	374	4.2	348	4	US-09-557-034-28	Sequence 28, Appl
17	370.5	4.2	949	3	US-09-196-392-10	Sequence 10, Appl
18	370.5	4.2	949	4	US-09-841-819-10	Sequence 10, Appl
19	363	4.1	1166	4	US-09-392-082-5	Sequence 5, Appli
20	356.5	4.0	352	3	US-09-065-474-139	Sequence 139, App
21	356.5	4.0	352	4	US-09-557-034-139	Sequence 139, App
22	353	4.0	302	2	US-09-031-495-38	Sequence 38, Appl
23	353	4.0	302	2	US-08-847-429A-38	Sequence 38, Appl
24	353	4.0	302	3	US-09-065-474-38	Sequence 38, Appl
25	353	4.0	302	4	US-09-557-034-38	Sequence 38, Appl
26	353	4.0	303	2	US-09-031-495-23	Sequence 23, Appl
27	353	4.0	303	2	US-08-847-429-23	Sequence 23, Appl
28	353	4.0	303	3	US-09-065-474-23	Sequence 23, Appl
29	353	4.0	303	4	US-09-557-034-23	Sequence 23, Appl
30	353	4.0	741	2	US-08-436-771-4	Sequence 4, Appli
31	353	4.0	741	2	US-08-434-998-4	Sequence 4, Appli
32	353	4.0	741	2	US-08-487-797-4	Sequence 4, Appli
33	353	4.0	741	5	PCT-US96-02058-4	Sequence 4, Appli
34	348	3.9	679	2	US-08-462-481-4	Sequence 4, Appli
35	348	3.9	679	2	US-08-436-771-6	Sequence 6, Appli
36	348	3.9	679	2	US-08-434-998-6	Sequence 6, Appli
37	348	3.9	679	2	US-08-487-797-6	Sequence 6, Appli
38	348	3.9	679	2	US-08-701-005A-4	Sequence 4, Appli
39	348	3.9	679	2	US-08-479-895-4	Sequence 4, Appli
40	348	3.9	679	5	PCT-US96-02058-6	Sequence 6, Appli
41	344	3.9	679	3	US-08-943-956A-4	Sequence 4, Appli
42	344	3.9	786	4	US-09-509-802-2	Sequence 2, Appli
43	344	3.9	787	3	US-09-188-930C-334	Sequence 334, App
44	344	3.9	787	4	US-09-312-283C-334	Sequence 334, App
45	343	3.9	741	2	US-08-462-481-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Gueglot, Karl C.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF.0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: gb0491
US-09-172-977-4

Query Match 6.0%; Score 530.5; DB 2; Length 1839;
Best Local Similarity 20.2%; Pred. No. 3.2e-36;
Matches 436; Conservative 263; Mismatches 654; Indels 763; Gaps 81;

Qy 14 EEENIPAMKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKNGANCNLEDLDNWT 73
Db 73 KEGHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFT 132

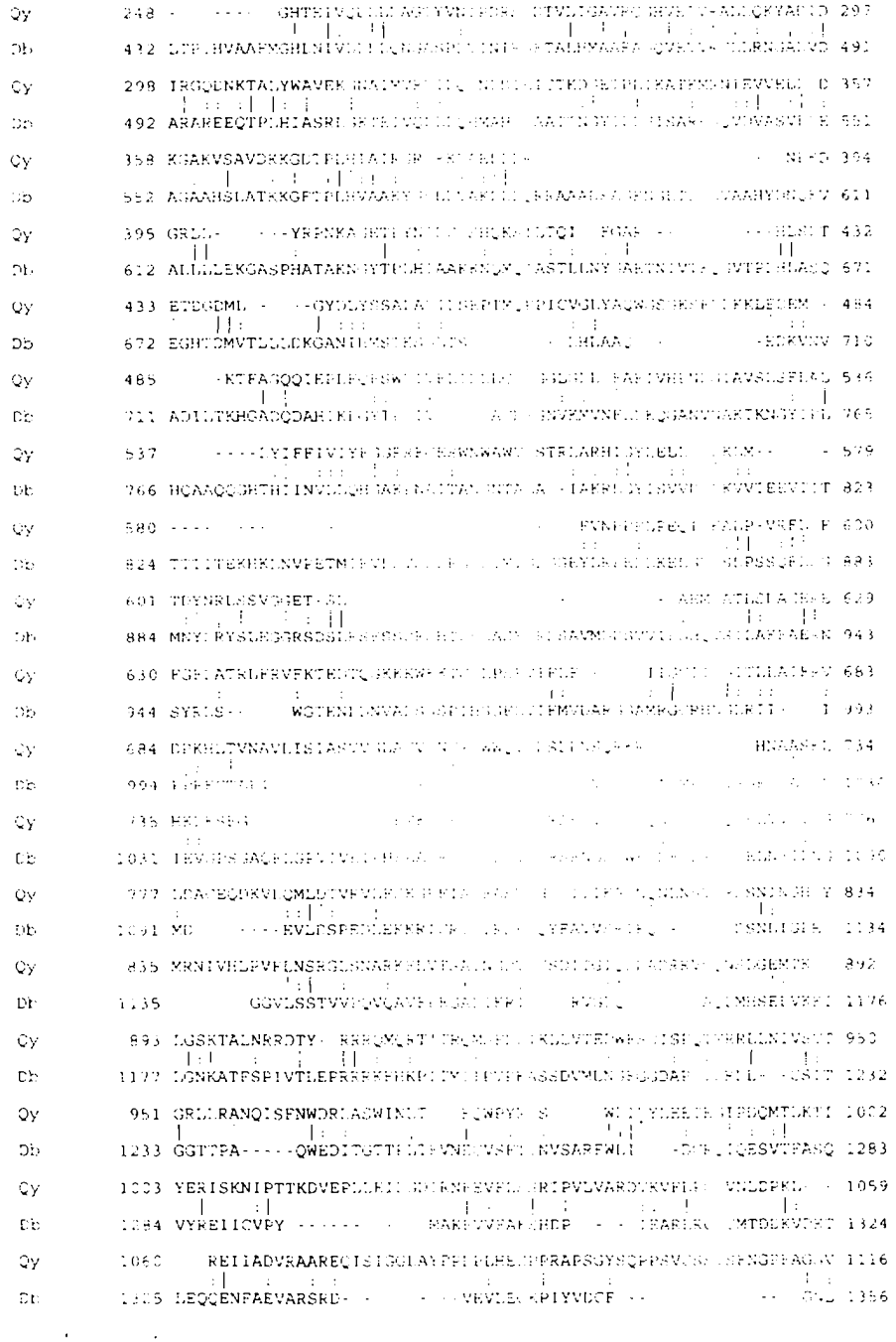
Qy 74 ALISASKEGHVHIVEELLKCGVN---------- LEHRDMG--- 103
Db 133 PLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVR 192

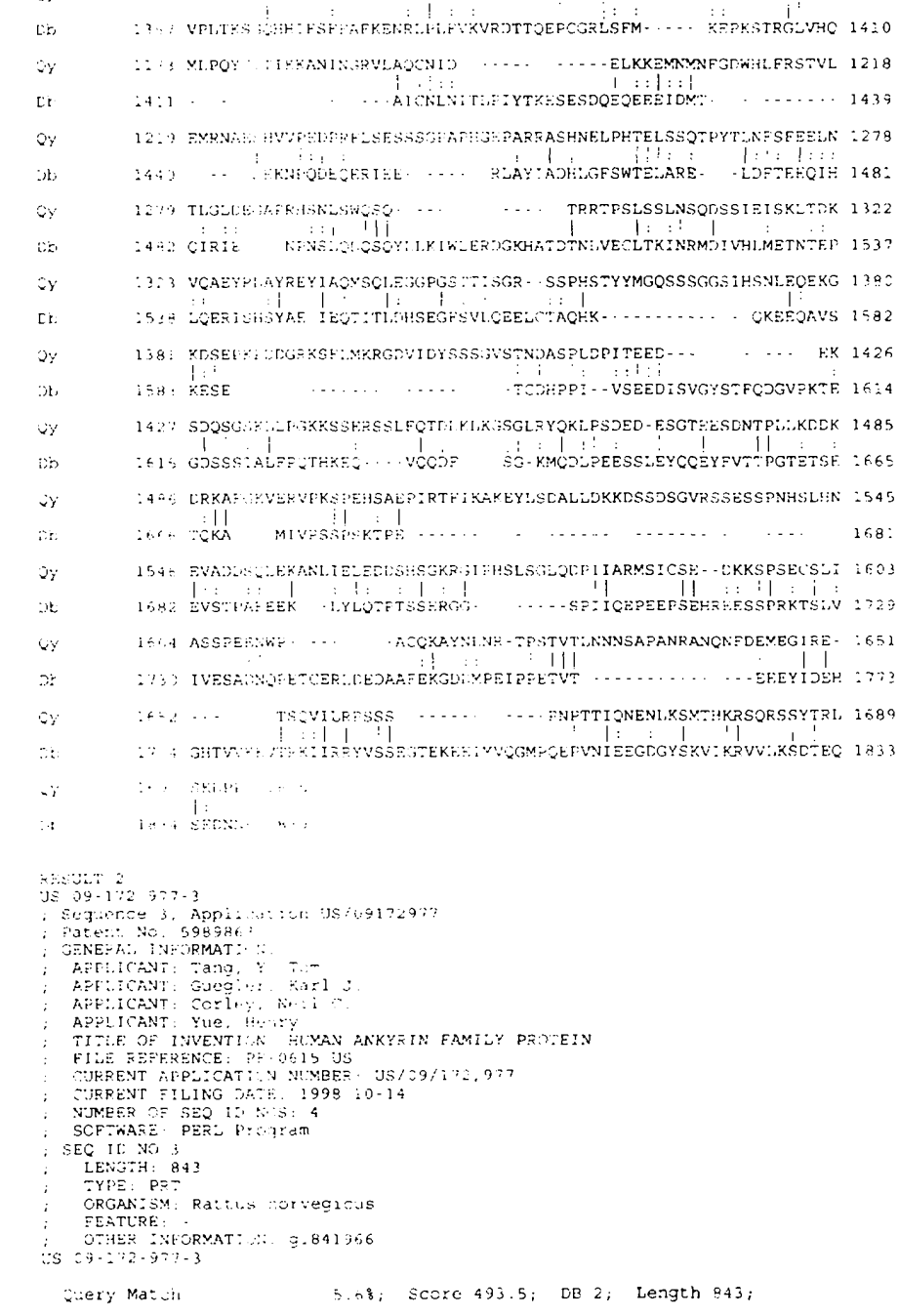
Qy 104 ----- ---------- ---- ----------GWTALMWACYKGRTDVVELL 123
Db 193 LPALHIAARKDDTKSAALLLQNDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLL 252

Qy 124 LSHGANPNVTDQLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKG- 182
Db 253 LNRGAAVQFTA-RNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 311

Qy 183 ------- ------ ----------HLECVKHLLAMGADVDQEGANSMTALIV 210
Db 312 DQVVELLIERGAPLLARTKNGLSPLHMAAQGDHVECVKHLLQHKAPVDDVTLDYLTALHV 371

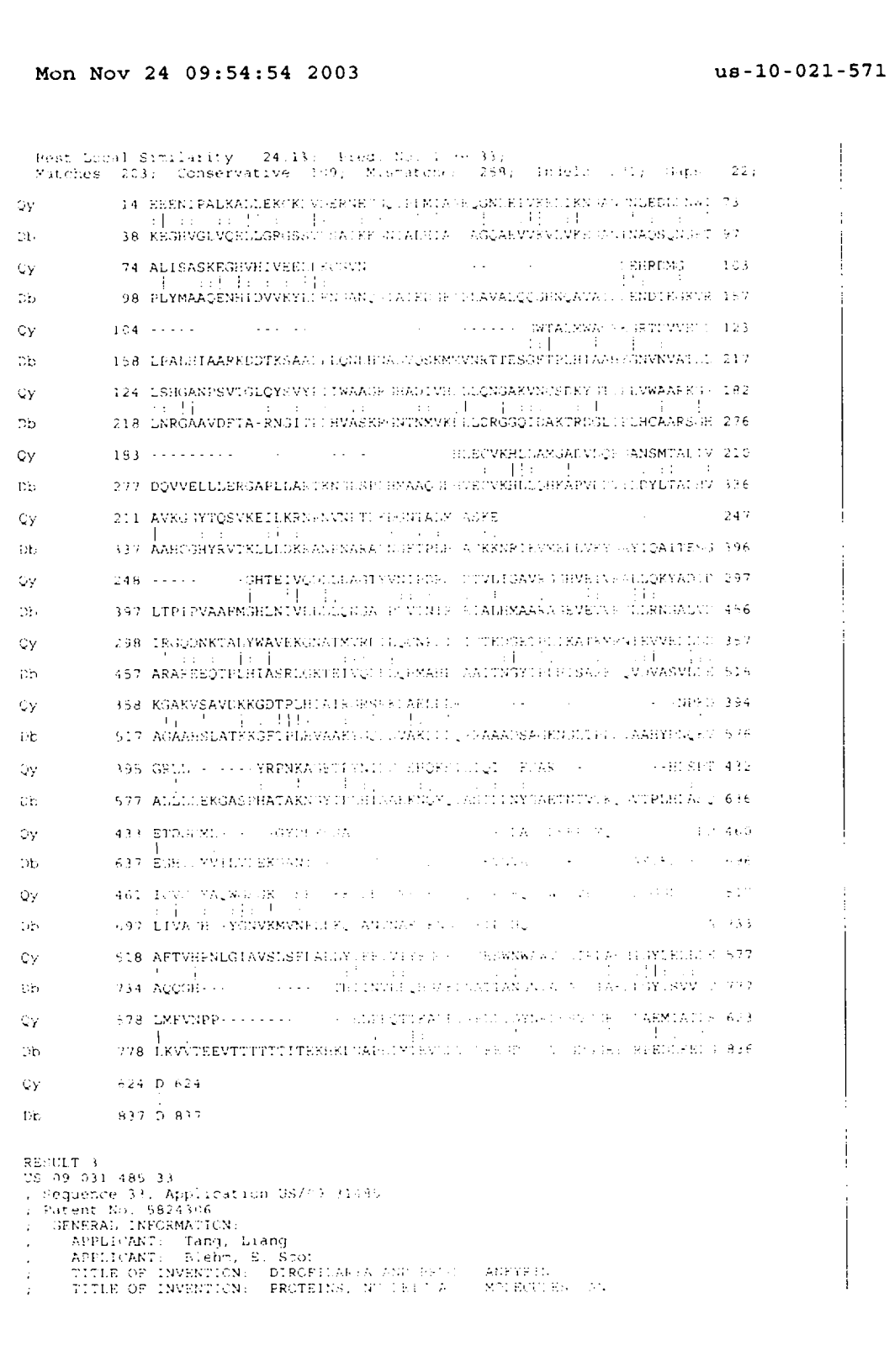
Qy 211 AVKGGYTDSVKEILKRNPNVNLTDKDGNTALMIASKE---------- ------ 247
Db 372 AAHCGHYRVTELLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGASIQAITESG 431





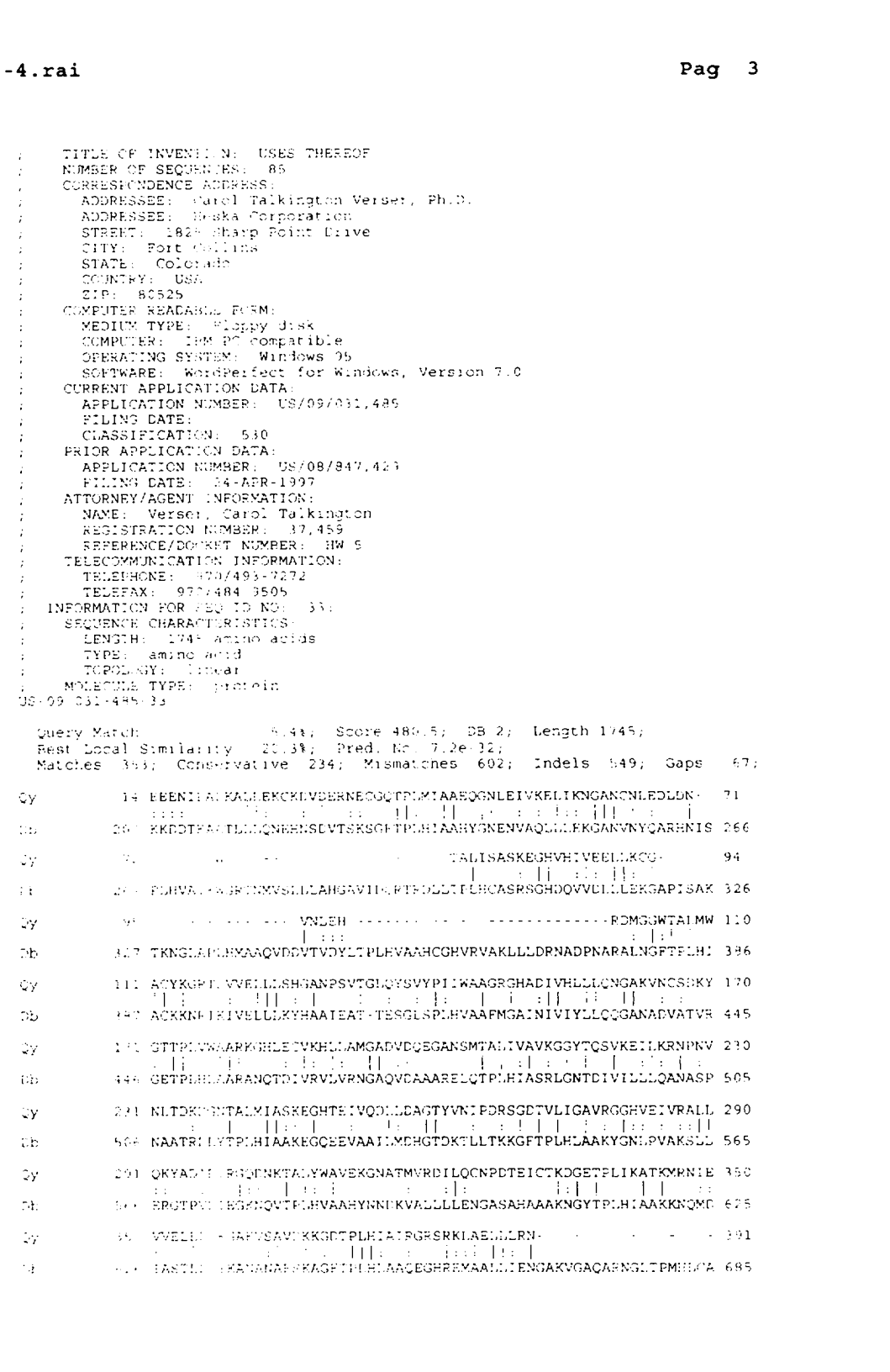
RESULT 2
US 09-172 977-3
; Sequence 3, Application US/09172977
; Patent No. 5989867
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Gueguler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: g.841966
US 09-172-977-3

Query Match 5.6%; Score 493.5; DB 2; Length 843;



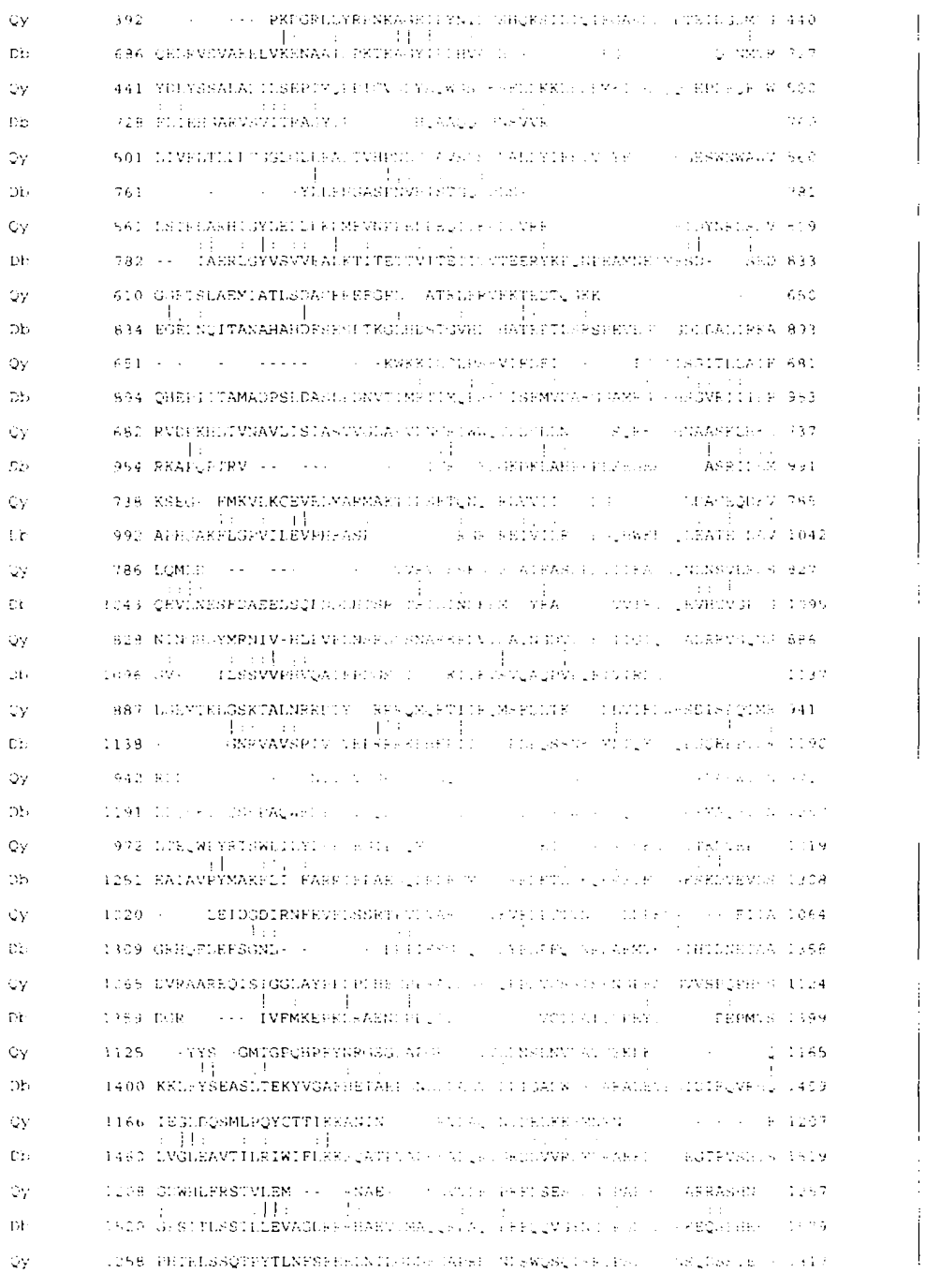
Best Local Similarity 24.1%; Pred. No. 1 e-33;
Matches 203; Conservative 109; Mismatches 259; Indels 17; Gaps 22;

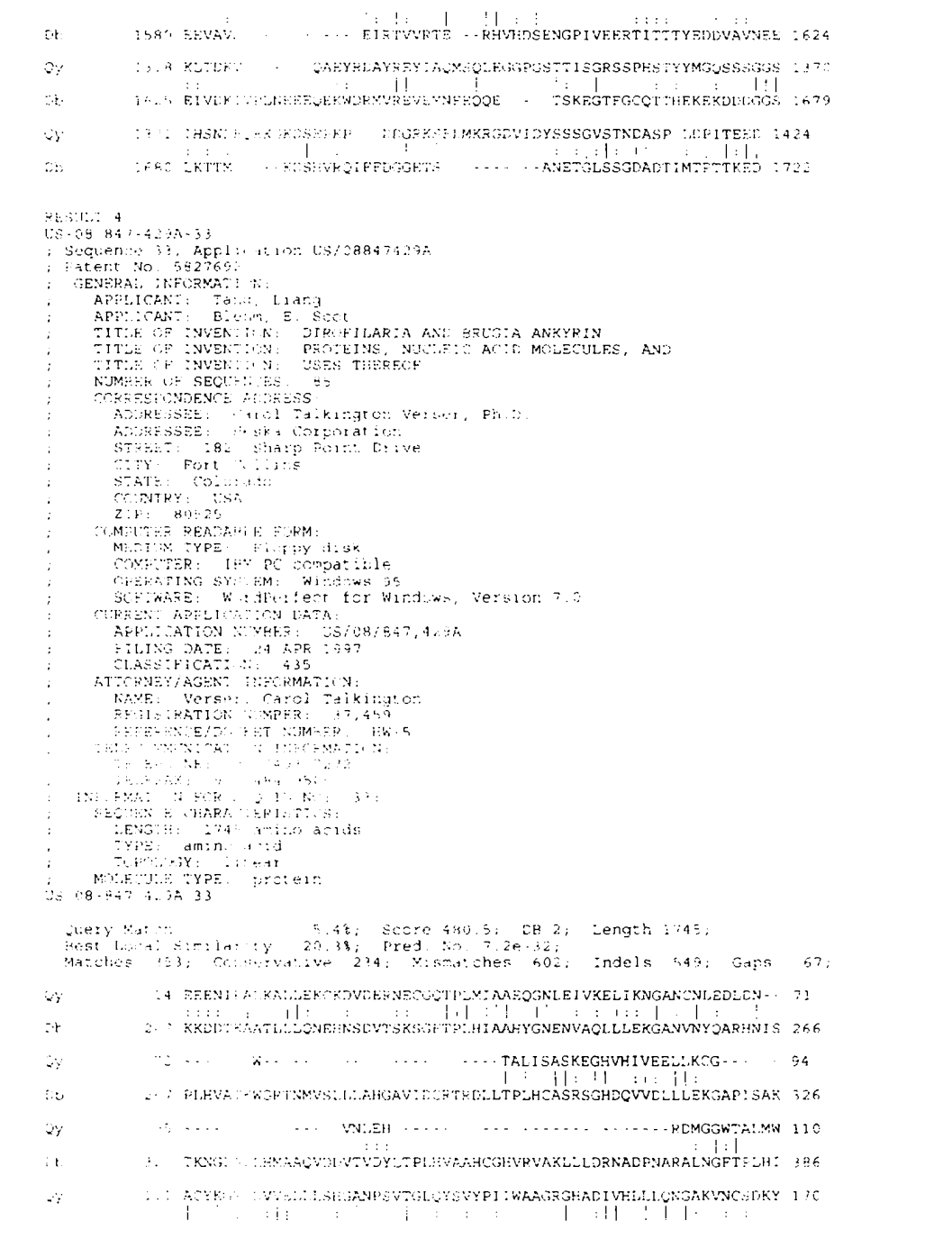
RESULT 3
US-09-031-485-33
; Sequence 33, Application US/09/031,485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Biehn, E. Scot
; TITLE OF INVENTION: DIHOFILARIA ANTIGENS
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND



; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,423
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW 5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484 3505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-33

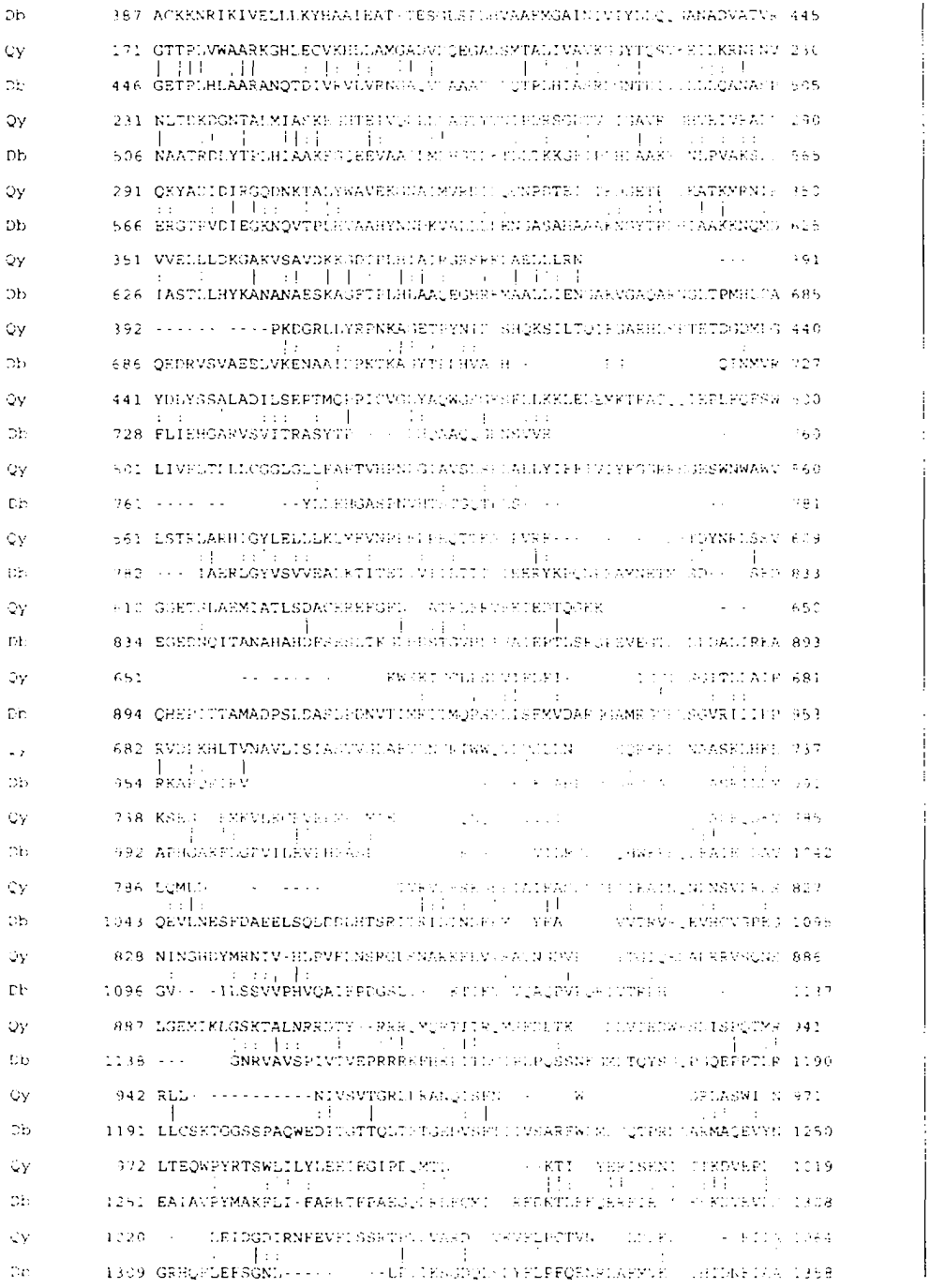
Query Match 5.4%; Score 480.5; DB 2; Length 1745;
Best Local Similarity 20.3%; Pred. No. 7.2e-32;
Matches 393; Conservative 234; Mismatches 602; Indels 549; Gaps 67;

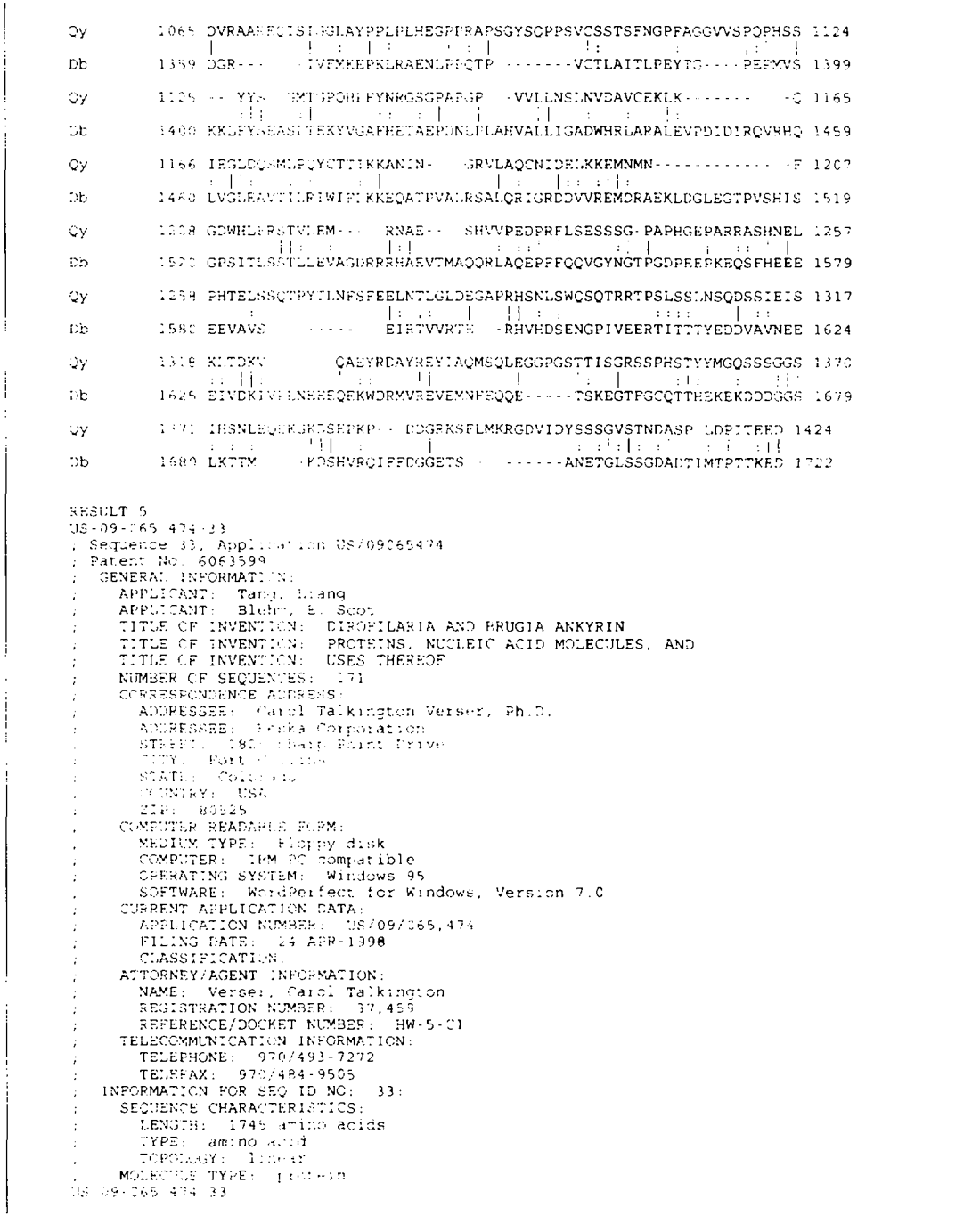




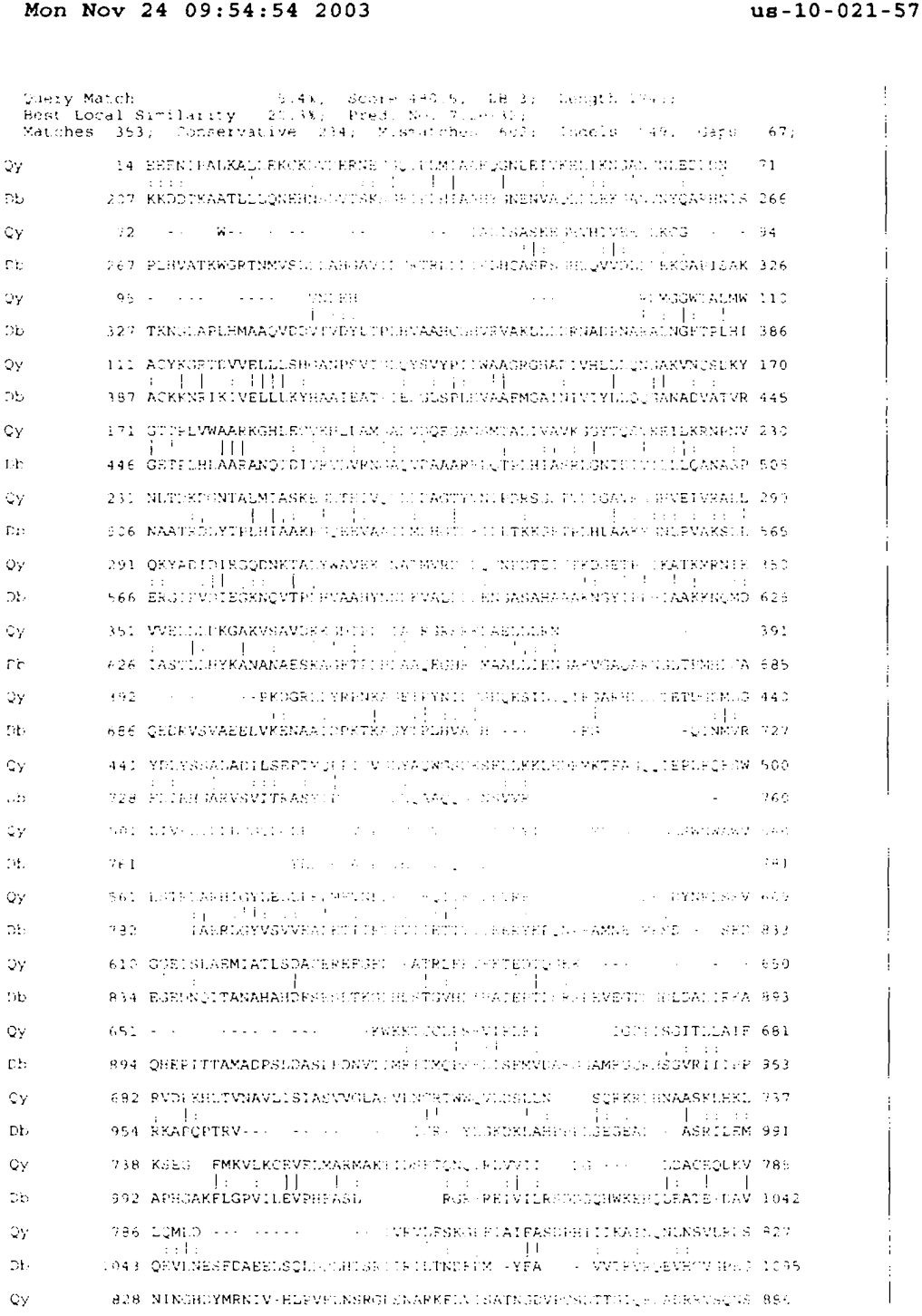
RESULT 4
US-08-847-439A-33
; Sequence 33, Application US/08847439A
; Patent No. 5827691
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Bloom, B. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Merck Corporation
; STREET: 181 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,439A
; FILING DATE: 24 APR 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: RW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (970) 226-5702
; TELEFAX: (970) 226-5611
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
Qy 08-847-439A-33

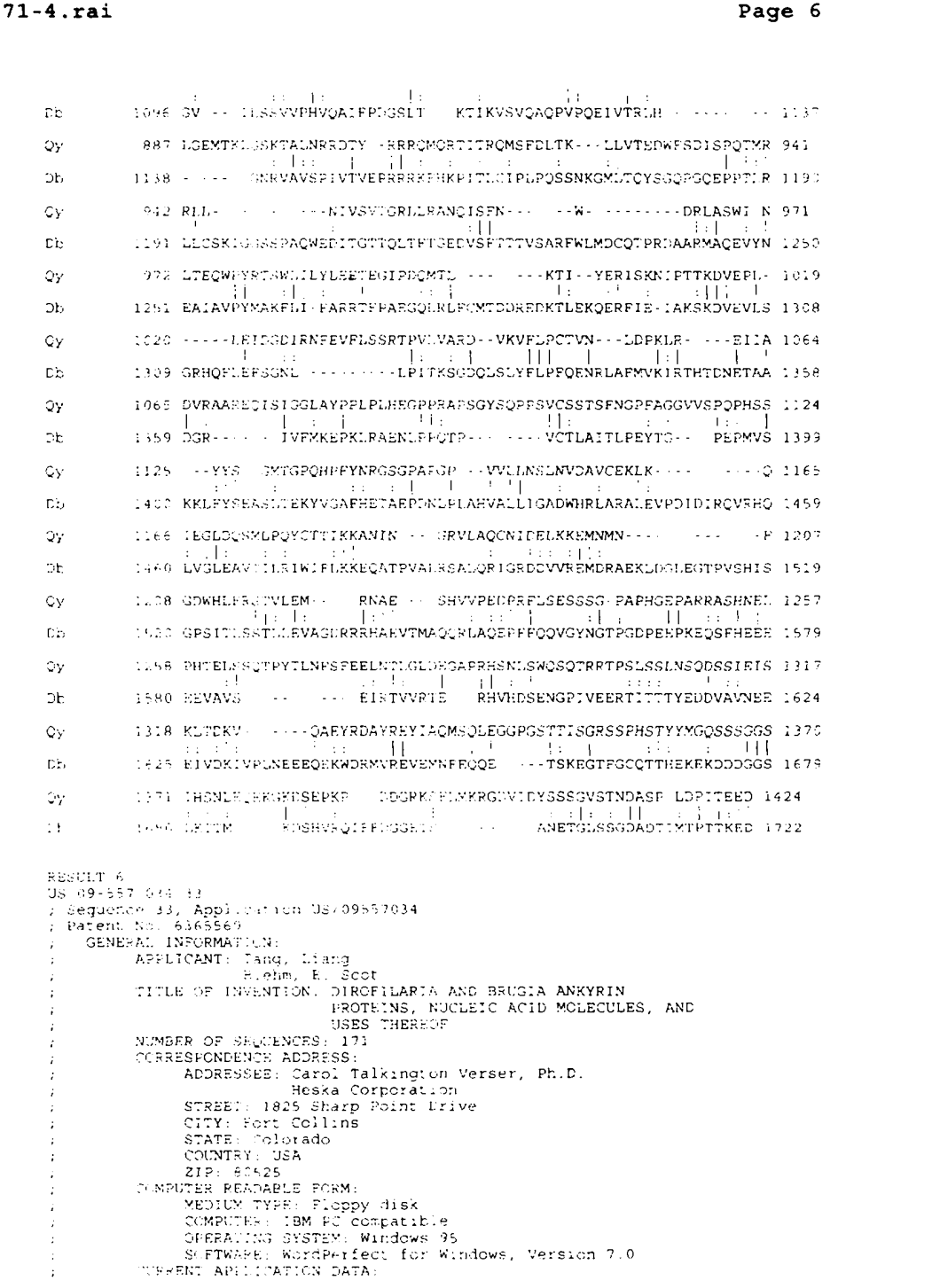
Query Match 5.4%; Score 440.5; DB 2; Length 1745;
Best Local Similarity 20.3%; Pred. No. 7.2e-32;
Matches 193; Conservative 234; Mismatches 602; Indels 549; Gaps 67;





RESULT 5
US-09-065,474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Leska Corporation
; STREET: 180 [illegible] Point Drive
; CITY: Port [illegible]
; STATE: Co[illegible]
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24 APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,456
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1746 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065,474-33

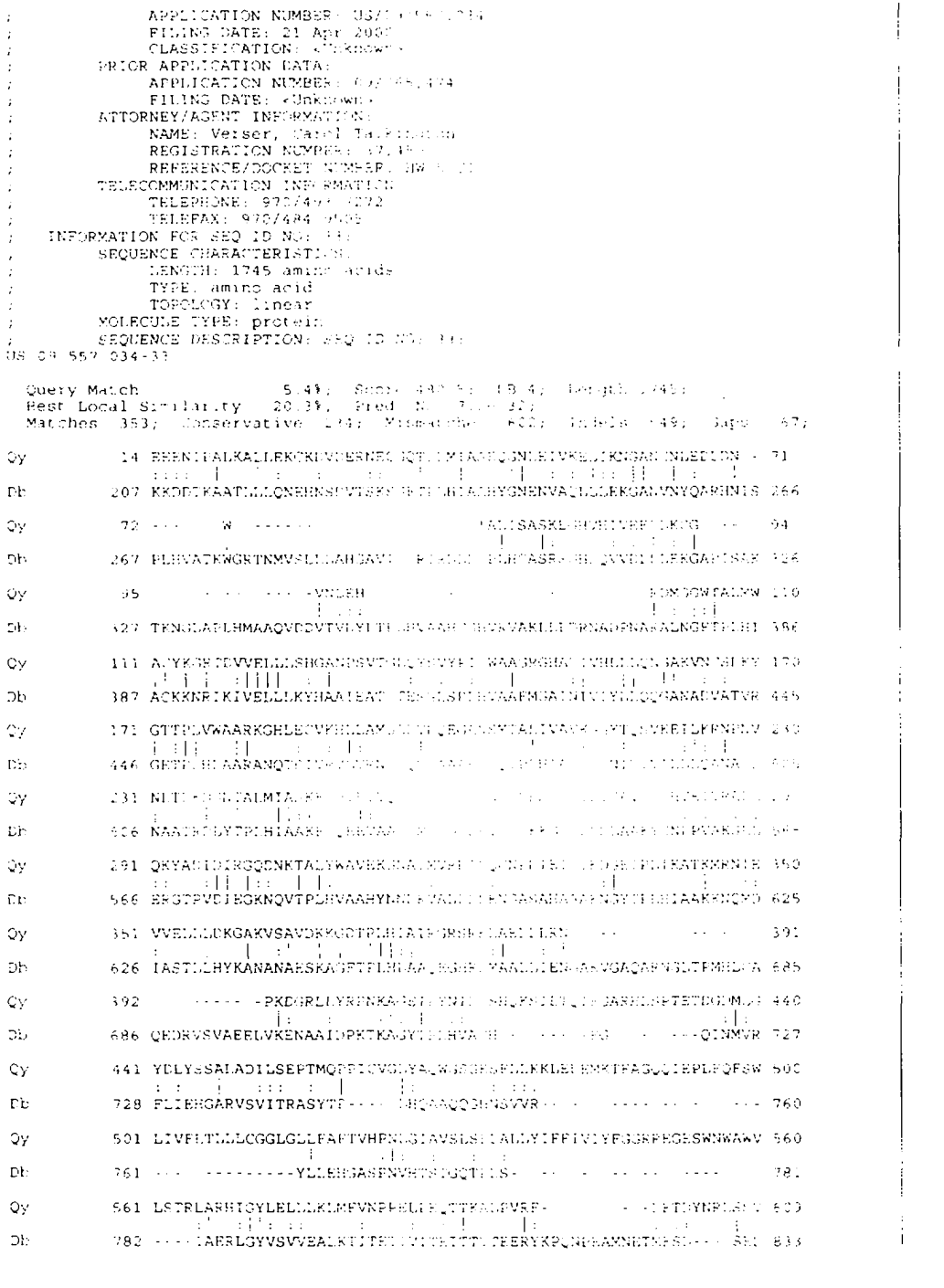




RESULT 6
US 09-557 946 33
; Sequence 33, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; Ehlm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/[illegible]
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Ta[illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/[illegible]
; TELEFAX: 970/484-[illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: [illegible]
US-09-552-934-33

Query Match 5.4%; Score [illegible]; DB 4; Length [illegible];
Best Local Similarity 20.3%; Pred. No. [illegible];
Matches 353; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];



RESULT 7
US-09-082-059-2
; Sequence 2, Application: US/09082059A

; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S
; APPLICANT: Devarajan, Prasad
; TITLE OF INVENTION: No. 6225086 Novel Calpain Fragments and Methods for Their Identificat
; FILE REFERENCE: 44574-5002 D
; CURRENT APPLICATION NUMBER: US/09/052,059
; CURRENT FILING DATE: 1998-03-31
; EARLIER APPLICATION NUMBER: 60/042,[illegible]
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-052-059-2

Query Match 4.0%; Score 108; DB 4; Length 1088;
Best Local Similarity 30.5%; Pred. No. 8.2e-26;
Matches 120; Conservative 71; Mismatches 125; Indels 41; Gaps 3;

Qy 14 EEEDIPALKALLEFRDVDRNECGQTPLMIAAEQGNLEIVKELIKNGA 62
Db 30 KENRIKVMELLLKHGASIQAVTE[illegible] 93

Qy 74 ALISASKEGHVHIVELLKCGVNLEHRDMGGWTALMWACYKGRTDVVEL 122
Db 90 ALHVAARSGQAEVVRYLIQNA[illegible] 143

Qy 123 TLSHGANISVTSLQYSVYFIIWAASRGRADIVRLLLQNGAKVNCSDKYGTTPLVWAARKG 182
Db 144 TLSNYT-PLHLSAFEGKVAIEIMVKT[illegible]VDAARSAGHLEVAKVLIQ 197

Qy 183 HLECVRLLAMGALVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTALM 242
Db 198 MIALVDQEGANSMTALIVAVEIAGIELVKALLEHGADVNAQDKFGKTAFD 252

Qy 243 IASKFGHTEIVQLLDAGTYVNIPDFSGRTVLIGAVRGGHVEIVRALLQKYADIDIRGQD 302
Db 253 VQIILDAGTYVNIPDF-TGVLI-VV-IHGGFTRALQEAIQSRLGFTALYWAVE 312

Qy 303 NKTALYWAVEKGNATMVRDILQQNPTEICTKDGETPLIKATKMRNIEVVELLLDKGAKV 362
Db 313 ATTIEYGADANAVTE[illegible] 367

Qy 363 SAVDKFGCTPLHIAIRGRSRKLAEILFNPKDDGKLQYRPNKAGETPYNIDCSHQKSILTQ 422
Db 368 EALTGQTPLAIAEQLARSEQHEHVAGLLAPLRKD---------- 408

Qy 423 IFGARHLDPIRTDGEMLGYELYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKFLED 482
Db 409 ISPTQ----------NLQPRIKLKLFGHSGSGKTTLVESLK 451

Qy 483 EMKTPAIQIESFLQPSWLIVFLTLLIGGGLGLLFAFTVHPNLG---------- 526
Db 452 ----------QPLLRSFFR-RRRPRLSSTNSSRFPPSPLASKP 489

Qy 527 IAVAIIILALLYIFFIVIYFGGRFDSWNAWVLSTRLARHIGYLEIIKLMFVNPFE 585
Db 490 TVSVI[illegible]YPGPFTVLGRSPGMMFEPGLTK--GMLE--VFVAPTH 535

Qy 586 [illegible] 619
Db 536 HPHGSI[illegible]KTAIIQNAYLNGVLPGVWEPSGNFVYFGGVDY---FAANDPTSIHVVV 592

Qy 620 ATLSLAIMSFFS[illegible]FIATFLFRVFFTEDT---QGKKKWKKTCCLPSFVIFLFIIGGIIS 673
Db 593 FSLEEPYEIQLNPVIPWISFLKSLVPVEEPIAPGGKLK------NPLQVVL-------- 637

Qy 674 GITLLAIFVDFKHLTVNAVLISIAVVGLAFVLNCRTWWQVLDSLLNSQRKRLHN---A 730
Db 638 ----------VATHADIMNVFPASGEFGYDKDT------SLLKEIRNRFGNDLHI 687

Qy 731 ASKLHFPLVSFGF MKVLKCEVELMARMAKTIDSFTQNQTRLVVIIDGLDACEQDKVL 786
Db 688 SNKLEVLLAVAGGGKDMKVLRNHIQ EIRSQIVSVCPPMTHL----- --CE--KII 963

Qy 787 QMLDIVRVLESKGPFIAI FASDPHIIIKAINQNLNSVLRDSNINGHDYMRNIVHLPVP 844
Db 964 STLPSWREMLNGPNQLMSLQQFVYD- VQDQLNPLASEED------LRRIAQQ--- 1007

Qy 845 LNSRGLSNARKFLVTSATNGDVPCSTTGIQEDADRRVSQNSLGEMTKLGSKTALNR-RD 903
Db 1008 LHSTGEIV IMQSETVQDVLLLD-------PRWLCTNVLGKLLSVETPRALHHYRG 1059

Qy 904 TYRRGMLPIIIRQMSPDLTKLLVTHDWFS-DISPQTMRRLLNIVSVTGRLLPANQISFN 962
Db 1060 RYTVEIIPVPDSDVEELLQILDANDICARDLSSGTM- ----VDVPA LIKTDNLHRS 1106

Qy 963 W 963

RESULT 8
US-08-810-712-10
; Sequence 10, Application US/08810712
; Patent No. 6161036
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co.
; TITLE OF INVENTION: Tumor Suppressor Protein, Pharmaceutical Composition and Diagnostic Kit
; TITLE OF INVENTION: Use of Said Genes and Proteins
; FILE REFERENCE: sequencelist
; CURRENT APPLICATION NUMBER: US/08/810,712
; CURRENT FILING DATE: 1997-03-03
; EARLIER APPLICATION NUMBER: PCT/US94/10[illegible]
; EARLIER FILING DATE: 1994-10-11
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1424
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-810-712-10

Query Match 4.0%; Score 108; DB 18; Length 1424;

```
Db        1109 W 1109

RESULT 9
US-09-392-812A-2
; Sequence 2, Application US/[illegible]
; Patent No. 6537778
; GENERAL INFORMATION:
;  APPLICANT:  Zuker, Charles S
;  APPLICANT:  Walker, Richard G
;  APPLICANT:  Willingham, Aaron
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION:  A Eukaryotic Membrane Protein Family
;  FILE REFERENCE:  9230/B-6976/GDG
;  CURRENT APPLICATION NUMBER: US/09/392,812A
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1704
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence for novel human receptor
;   OTHER INFORMATION: sequence
US-09-392-812A-2

  Query Match              4.3%;  Score 174.5;  DB 4;  Length 1704;
  Best Local Similarity   22.4%;  Pred. No. 5.6e-23;
  Matches  169;  Conservative  104;  Mismatches  241;  Indels  [illegible];  Gaps  39;

Qy          29 KDVSGRNESGQTPLMIAAQQGNIEIVKELIKNGANCNLEDLDNWTPLHLAAGYNRVRIVQLLQHGADVHAKDKGGLV 108
               |  :: :  | |:::|  : :::  .  :  .  :  .  |  :  :  :|:  :
Db        516 KAMNRDSVGNTPLLIAMRHPEIVDLLIANGARVDVRDTFGRSALAAAVYLRVEIVDVLLKYGAEPNVPDNFGITPLHM 595

Qy          89 ELLRNGVNLEHRDMQGWTALMWACSYGHIEIVEFLIQHGVSLDVQNKFGETALDIARRAGQTEIVTLLKSAG 162
               ||   :  :  :|:  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        596 ALITNRAPINSKSRVSPTALHLAAMNGHTHIVKILIKHNADVNIQDKFLQTPLHLAAA 654

Qy         149 RGHADIVRLLLQNGAKVDSRDKYGITAWAAAKGHLEVVKLLLANGSAMNRIA 200
               :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         655 SGLMEVTQLLLEIGANIFATDKLPRTPLHLAVEHGHDKVVSLLDKYNVDVNAQ 708

Qy         201 LIVANGQDYTLQVKEIPGGVTLEMFDVKTNPARTETVTKQPSRLASFTVHQFCFIYNPDYEFVQRPLDLTGIMIDPKN 280
               :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         709 LIVANGQDYTLQVKEIPGGVTLEMFDVKTNPARTETVTKQPSRLASFTVHQFCFIYNPDYEFVQRPLDLTGIMIDPKN 788

Qy         281 DRGHCISARNKLTDATSLAAPGKAGLYRSLFMSEQCSKHRGKTSRKSSKGGLAAQDHLSSPLVPSESDDHVLSNSSL 360
               :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         789 DRGHCISARNKLTDATSLAAPGKAGLYRSLFMSEQCSKHRGKTSRKSSKGGLAAQDHLSSPLVPSESDDHVLSNSSL 868

Qy         361 --IDIRGQDNK---  ----TALHWACEKGDVVMVEVLLKKGAKVNAVNKWGDTPLDEAVYFGHEELVKLLESHGGHS 430
                 :  :  :|:   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         869 QVLDVLKSTNSLRINSKKLSDIPLVAAYYGLVDVLQSLIDAGVDVNAKTSQGWTPLDEAVYFGHEELVKLLESHGGHS 948

Qy         431 ICIKHGETPL------  ----------  ----------  ----------  ----------  ----------  ---- 464
               :  :  :  :|:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         949 LGTESGMTPLHLAAFSGNENVVRILVAAGCQLNAATKDGMTPLQVAVFFGQGAFDNA 1008

Qy         465 VDKKHDTPLHIAIRGRSRKLAELLL  NIA  ----------  ----------  LLYFPNE 502
               :  :  :|:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1009 TDRNGWTPLHCAAKAGHLEVVKLLLQAGASIDAVTFAAIWAAIEGHLDLRYLMNF 1064

Qy         503 AGET-  ------PYN  IDCSHLPSFI  LDPSHLSIAI  DMLESDTYSSALAIS 551
                  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1065 EHDTYSLMEDKRPVYNLMVVSPKDGDNFTPLHCASSRAIVAKRLCLYGASTPFEA 1123

Qy         552 LSEPTMQPPICVGLYAQW  ----------  ----------  ----------  ----------  ---- 599
               :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1124 AKDLVAAGKQCEAMATELIALAAKTKQKSLLDVSHVDEDTLLKLAVQLGVEDIAQTV 1183

Qy         600 GQQTEPLPQ  ----FSNLIVFLTL  ----TGLTAPTVHEV  ----------  ----  ---- 649
               :  :  :|:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        1184 GKDTEPLPQ  ----FSNLIVFLTL  ----TGLTAPTVHEV  ----------  ----  ---- 1233

Db        1234 QRYLQSIWRGSLTWASWKILL  ----  -LLVAFIVCPPVWIGFTFPM----------  913

Qy         644 IYFGGFREGRSWNWAWVLS--TPLAHIIGYLELLLKLMFVNPPELPEQTTKALPVRFLFT 701
                  |  :  :  :  :  :  :|  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         914  -  QHSFNKVPTIKFMSYLTSHI-YLMIHLSIVGITP--  --  IYPVL---  951

Qy         702 DYNRLP  SVGGRTSIAEMIAIISDACEREPGFLATRLFRVFKTEDTQGKKKWKKT 755
                  :|,  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         952 ---PLSLV.YWYEVS  -  ---LLIWLG  -----GLL--LFELTNPSDKSG---  --  984

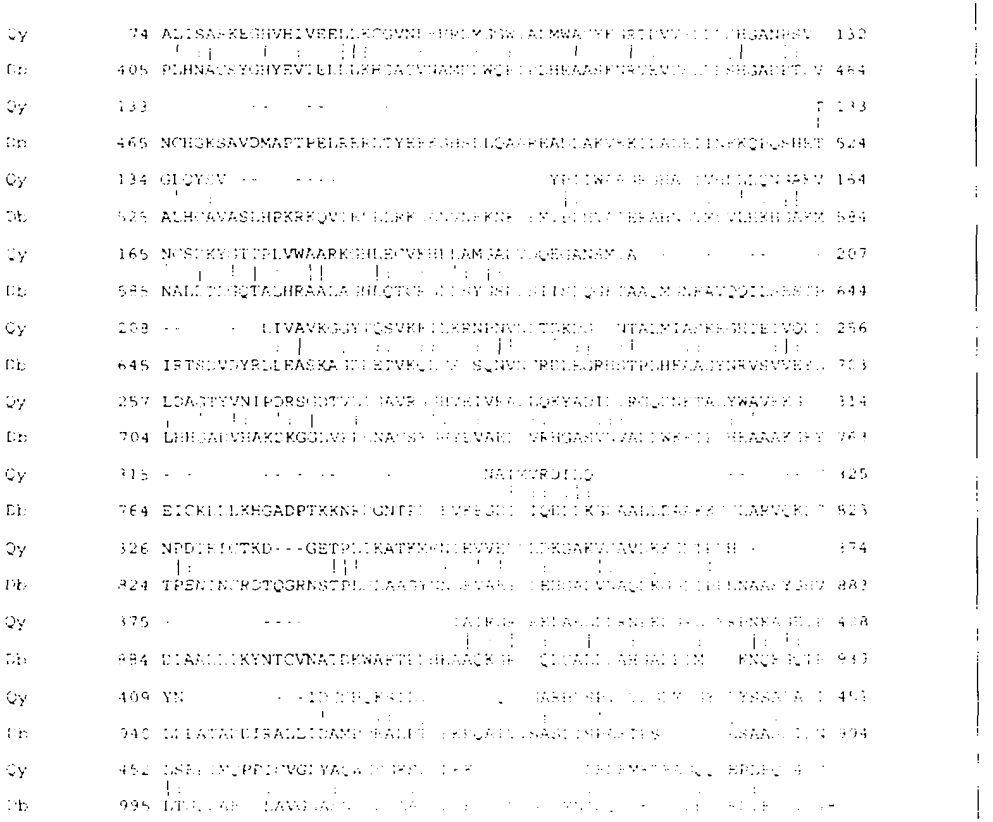
Qy         756 CCLPSPVIHLFIIGCIISGIILIAIFPVDPKHLTVNAVLISIASVVGLAFVLNCRTWWQV 815
                  :  :  :  :|:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         985   LGSLRVLITLLRMAIVGVFVSAFLVSKEYWP--TLVYGRNQCFALAFLLAC  -VQI 1037

Qy         816 LDSL 819
                  |
Db        1038 LDFL 1041


RESULT 10
US-09-196-387-2
; Sequence 2, Application US/09196387
; Patent No. 6277611
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1 [illegible], Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/995,105
;      FILING DATE:  [illegible] 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-240 PIPI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133503
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-196-387-2

  Query Match              4.3%;  Score 174.5;  DB 3;  Length 1327;
  Best Local Similarity   22.4%;  Pred. No. 5.6e-23;
  Matches  169;  Conservative  104;  Mismatches  241;  Indels  247;  Gaps  39;

Qy          14 EEENIRALFALLERKDVDFRNETGQTPLMIAARQGNLEIVKELIKNGANCNLEDLDNWT 73
                  |:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         948 EEFLNALIPLDVNCHAEDGFK-  N.PLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 1004
```



RESULT 11
US 09/841,846-2
; Sequence 2, Application US
; Patent No. 6506...
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN
; TITLE OF INVENTION: OF USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 09/841,846
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson, Esq., David A.
; REGISTRATION NUMBER: 36,779
; REFERENCE/DOCKET NUMBER: 600-1-247 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5900
; TELEFAX: 201-343-1684
; TELEX: 133520
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US 09/841,846-2

Query Match 4.2%; Score 406.5; DB 4; Length 1048;
Best Local Similarity 22.3%; Pred. No. 5.6e-23;
Matches 164; Conservative 89; Mismatches 220; Indels 247; Gaps 17;



RESULT 12
US-09-392-812A-4
; Sequence 4, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S
; APPLICANT: Walker, Richard G
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 22307E-097600US
; CURRENT APPLICATION NUMBER: US/09/392,812A
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1619
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC cDNA
; OTHER INFORMATION: sequence
US-09-392-812A-4

Query Match 4.2%; Score 74; DB 4; Length 1619;
Best Local Similarity 25.1%; Pred. No. 8.1e-23;
Matches 136; Conservative 97; Mismatches 202; Indels 1; Gaps 18;

Qy 29 KDVDERNECGQTPLMIAARQGNIEIVKFLIENGANVNLEDIDNWTALIWAVKRGHVSIVE 88
Db 661 KAMNEQSSVGWTPLILACHRGHVEVVNVLLAHSPVDVFDTEHRSALHLAVERGYIDVVD 720

Qy 89 ELLKCGVNLEHRDMSGWTALMWA 147
Db 721 ALLTNKAFINSKSRVGRTALHLAAYN 769

Qy 148 RGHADIVHLLLQNGAKVNQSDKY 207
Db 770 SGQMEVCQLLLELGANIDATDLW 798

Qy 208 LIVAVKGGYTQSVKEILKRNPN 264
Db 799 VAAQNNYSEVAKLFLQQRPSIVAINSDGN 856

Qy 265 IPDRSGDTVLIGAVRGGHVE 284
Db 857 ARNFIGATPLQLAAGFGRAEIV 916

Qy 301 QDNK 305
Db 917 STNSLRINSKKDGLTPLHVAAYYGHAEIVRELIVRATVNPLTVQGLTPLGLTSSQM 976

Qy 306 TPLIKATKMRNIEVVELLDK-GAKVSAVD 333
Db 977 TPLHLAAFSGNENVVRLLLNSAGVLVAATEL-GYNPLHLACFAGHMSVVELLL-QRSA 1034

Qy 334 RLLYSPNKAGETPYNIDCSHQKSIKTLIP 362
Db 1035 ELLQSQDRNGRTGLHIAAMHGHIQMVEILGQ-HA-EINATDFNG---WTPLHCAAKAGHL 1090

Qy 450 ---DILSEPTMQPPI-----CVGLVAWSGKPLLKKLEDFYRIFAGQQTEPLFQFSWL 501
Db 1091 EVVKLLCEAGASPKSETNYGCAALWFAAKEGHNEVLRYLMDREHDTYGLMPLKRPVYNLM 1150

Qy 502 IV 503
Db 1151 VV 1152

RESULT 13
US-09-031-485-28
; Sequence 28, Application US/09031485
; Patent No. 5824806
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blaxter, M. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,420
; FILING DATE: 24 APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-28

Query Match 4.2%; Score 74; DB 2; Length 348;
Best Local Similarity 33.7%; Pred. No. 7e-24;
Matches 117; Conservative 48; Mismatches 144; Indels 38; Gaps 4;

Qy 45 AAEQGNIEIVKELIRNGANVNLEDLGWTALISASKEGHVHIVEELLKCGVNLEHRDMGG 104
Db 44 AANAGHLDIVDELLERSGTDINICNANGLNALHLASKEGHHEVVRELLKRKADVDAATRKG 103

Qy 105 WTALMWAVYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKV 164
Db 104 NTALHIASLAGQELIVTVLVENGANVNVQSLN-GFTPLYMAAQENHESVVRYLLAHNANQ 162

Qy 165 NQSDKYGITPLVWAAAKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEIL 224
Db 163 ALSTEDGFTPLAVALQQGHDRVVAVLL---ENDTRGKVRLPALHIAAKKDDTKAATLLL 218

Qy 225 KRNPNVNLTDKDGNTALMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVE 284
Db 219 QNEHNSDVTSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTN 278

Qy 285 IVRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKAT 344
Db 279 MVSLLIAHGAVIDCR------ ----------TRDLL----------TPLHCAS 305

Qy 345 KMRNIEVVELLDKGAKVSAVDKKGDTPLHIA 376
Db 306 RSGHDQVVDLLLEKGAPISAKTKNGLAPLHMA 337

RESULT 14
US-08-847,420A-28

```
; Sequence 28, Application US/08847424A
; Patent No. 5827691
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Bloom, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AN[illegible] A ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, N[illegible] AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,424A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5 CT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-424A-28

  Query Match              4.2%;  Score 274;  DB 2;  Length 348;
  Best Local Similarity   33.7%;  Pred. No. 7e-24;
  Matches  112;  Conservative   48;  Mismatches  134;  Indels   38;  Gaps    4;

Qy          45 AARQ[illegible]IVPEL[illegible]PDGAN[illegible]QLED[illegible]DW.ALISASKEGHVHIVEELLKQHVNLEHRDMGG 104
               :  :|| :  :  :  :  :  :  |  :|  :  |||:  :::  |
Db          44 AARA[illegible]AVELLR[illegible]GTDINTRNAL[illegible]NALHLASKEGHHEVVRELLKPKADVDAATRKG 103

Qy         105 WTALMWAAYKGRTDVVELLLSHGANIDVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKV 164
               |  |  |  :  :|  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         104 NTALHIASLAGQEDIVTVLVKNGAMVNVQSLN GFTPLYMAAQENHESVVRYLLAHNANQ 162

Qy         165 NCSDKYGTTPLVWAARKGHLECVKHLIAMGADVDQEGANSMTALIVAVKGGYTQSVKEIL 224
               |  :  ||  |  :  :||  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         163 ALSTRNGFTPLAVALQQGHDRVVAVLL  ENDIRGKVRLPALHIAAKKDDTKAATLLL 218

Qy         225 KRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVE 284
               :  :  :  :  |  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         219 QNEHNSDVTSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTN 278

Qy         285 IVRALLQKYADIDIPGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKAT 344
               :|  :  ||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         279 MVSLLLAHGAVIDCR---  ----  --  ----TPDLL--------  -- TPLHCAS 325

Qy         345 KMRNIEVVELLLDKGAKVSAVDKKGETPLHIA 376
               :  :  :  :|  :||  :||  :  |  :|:|
Db         326 RSGHEQVVDLLLEKGAPISAKTKNGLAPLHMA 357


Search completed: November 21, 2003, 12:48:59
Job time : 37 secs
```

```
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Bloom, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AN[illegible] A ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5 CT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-065-474-28

  Query Match              4.2%;  Score 274;  DB 3;  Length 348;
  Best Local Similarity   33.7%;  Pred. No. 7e-24;
  Matches  112;  Conservative   48;  Mismatches  134;  Indels   38;  Gaps    4;

Qy          45 AARQ[illegible]IVPEL[illegible]PDGAN[illegible]QLED[illegible]DW.ALISASKEGHVHIVEELLKQHVNLEHRDMGG 104
               :  :|| :  :  :  :  :  :  |  :|  :  |||:  :::  |
Db          44 AARA[illegible]AVELLR[illegible]GTDINTRNAL[illegible]NALHLASKEGHHEVVRELLKPKADVDAATRKG 103

Qy         105 WTALMWAAYKGRTDVVELLLSHGANIDVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKV 164
               |  |  |  :  :|  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         104 NTALHIASLAGQEDIVTVLVKNGAMVNVQSLN GFTPLYMAAQENHESVVRYLLAHNANQ 162

Qy         165 NCSDKYGTTPLVWAARKGHLECVKHLIAMGADVDQEGANSMTALIVAVKGGYTQSVKEIL 224
               |  :  ||  |  :  :||  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         163 ALSTRNGFTPLAVALQQGHDRVVAVLL  ENDIRGKVRLPALHIAAKKDDTKAATLLL 218

Qy         225 KRNPNVNLTDKDGNTALMIASKEGHIEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVE 284
               :  :  :  :  |  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         219 QNEHNSDVTSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTN 278

Qy         285 IVRALLQKYADIDIPGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKAT 344
               :|  :  ||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db         279 MVSLLLAHGAVIDCR---  ----  --  ----TPDLL--------  -- TPLHCAS 325

Qy         345 KMRNIEVVELLLDKGAKVSAVDKKGETPLHIA 376
               :  :  :  :|  :||  :||  :  |  :|:|
Db         326 RSGHEQVVDLLLEKGAPISAKTKNGLAPLHMA 357


Search completed: November 21, 2003, 12:48:59
Job time : 37 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

GM protein - protein search, using sw model

Run on: November 21, 20[illegible], 12:[illegible] ; Search time [illegible] Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 8864
Sequence: 1 MSVLISQSVINYVEEENIPA[illegible]HAAASSH GPSSPPAAL 1031

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96178682 residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distributions.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5205.5	58.6	1031	2	T43458	hypothetical protein DKFZp434F0621.1 - human
2	1742.5	19.7	1424	1	S22454	hypothetical protein - rat
3	1256	14.2	4454	2	T19552	hypothetical protein - mouse
4	555.5	6.3	1881	2	S38828	ankyrin 1, erythroc
5	542.5	6.1	1856	2	B35504	ankyrin 1, erythroc
6	532.5	6.0	1880	2	A35504	ankyrin 1, erythroc
7	544	6.1	426	2	AB2147	hypothetical protein
8	539	6.1	3924	2	S37431	ankyrin 2, neuronal
9	530	6.0	1961	2	T42716	ankyrin 3, splice
10	529.5	6.0	1940	2	T42719	ankyrin 3, splice
11	528	5.9	1862	2	I49561	ankyrin - mouse
12	523	5.9	1943	2	T42711	ankyrin 3, splice
13	521	5.9	4377	2	A55629	ankyrin 3, long sp
14	516	5.8	1765	2	T42713	ankyrin 3, splice
15	506	5.7	1848	2	S3777[illegible]	ankyrin, erythrocy
16	485.5	5.4	1549	2	T13941	ankyrin - fruit fl
17	442	5.0	2039	2	T19634	ankyrin related un
18	415	4.7	2584	2	T24159	hypothetical prote
19	415	4.7	2626	2	T24157	hypothetical prote
20	409.5	4.6	791	2	T42691	hypothetical prote
21	409.5	4.6	1423	1	I37227	death associated p
22	397.5	4.4	1062	2	T30257	invertebrin - mouse
23	375.5	4.2	323	2	B47164	ankyrin-like repea
24	373.5	4.2	557	2	T46502	hypothetical prote
25	373.5	4.2	1062	2	T14282	low protein - mous
26	349.5	3.9	397	2	T46447	hypothetical protein
27	348	3.9	679	2	B45577	2.5A tRNA-dependent
28	343	3.9	741	2	A46572	2.5A tRNA-dependent
29	336.5	3.8	991	2	T25411	hypothetical protein
30	329	3.7	1188	2	T19552	hypothetical prote
31	328	3.7	1401	2	S11527	alpha-latrotoxin p
32	325	3.7	934	1	B71274	probable ankyrin -
33	320.5	3.6	441	2	S74626	erythroid ankyrin
34	305	3.4	606	2	A42508	hypothetical prote
35	300.5	3.4	1411	2	S30356	alpha-latroinsecto
36	300.5	3.4	2004	2	D88948	protein ZK1005.1 [
37	300	3.4	1058	2	D82654	ankyrin-like prote
38	293.5	3.3	627	2	B86190	hypothetical prote
39	288.5	3.2	662	2	B84726	ankyrin-like prote
40	288	3.2	1184	2	T00253	gene Ankh2n protei
41	286	3.2	917	2	T48283	ankyrin-like prote
42	285	3.2	468	2	T18184	ankyrin repeat pro
43	285	3.2	633	2	T27499	hypothetical prote
44	284	3.2	1083	2	S46465	probable membrane
45	280	3.2	1001	2	S10385	G9a protein - huma

ALIGNMENTS

RESULT 1
T43458
hypothetical protein DKFZp434F0621.1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43458
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: A22561
A;Accession: T43458
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1031 <AAA>
A;Cross-references: EMBL:AL133620
A;Experimental source: adult testis; clone DKFZp434F0621
C;Genetics:
A;Note: DKFZp434F0621.1

Query Match 58.6%; Score 5205.5; DB 2; Length 1031;
Best Local Similarity 98.5%; Pred. No. 4.2e-287;
Matches 1004; Conservative 2; Mismatches 0; Indels 13; Gaps 1;

Qy 1 MSVLISQSVINYVEEENIPAAKALLEKCVDEPNRQSQTPLMIAAEQGNLEIVKELIKN 60
Db 1 MSVLISQSVINYVEEENIPAAKALLEKCVDEPNRQSQTPLMIAAEQGNLEIVKELIKN 60

Qy 61 GANCNIKDLDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120
Db 61 GANCNIKDLDNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVV 120

Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180

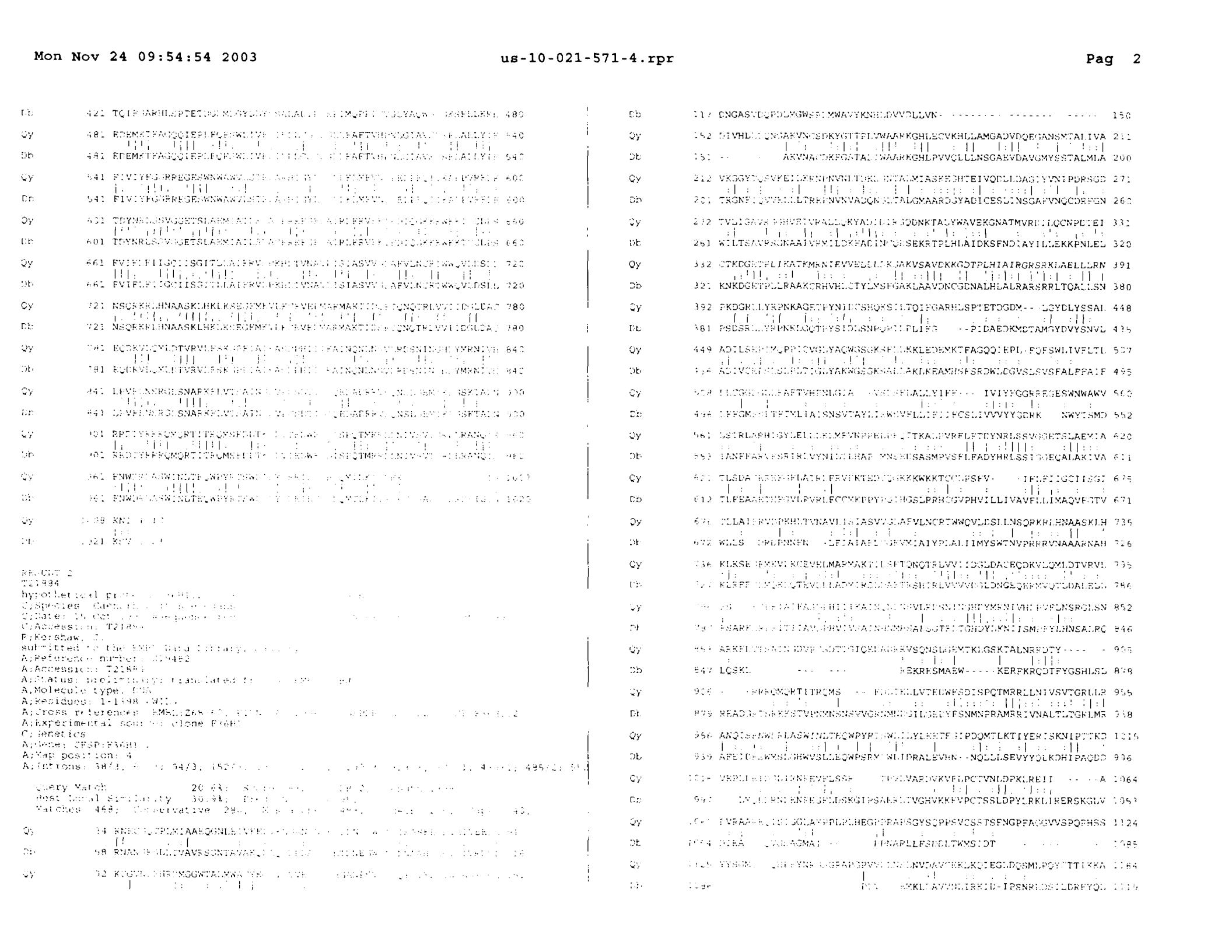
Qy 181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240
Db 181 KGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKDGNTA 240

Qy 241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
Db 241 LMIASKEGHTEIVQDLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Qy 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360
Db 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLLDKGA 360

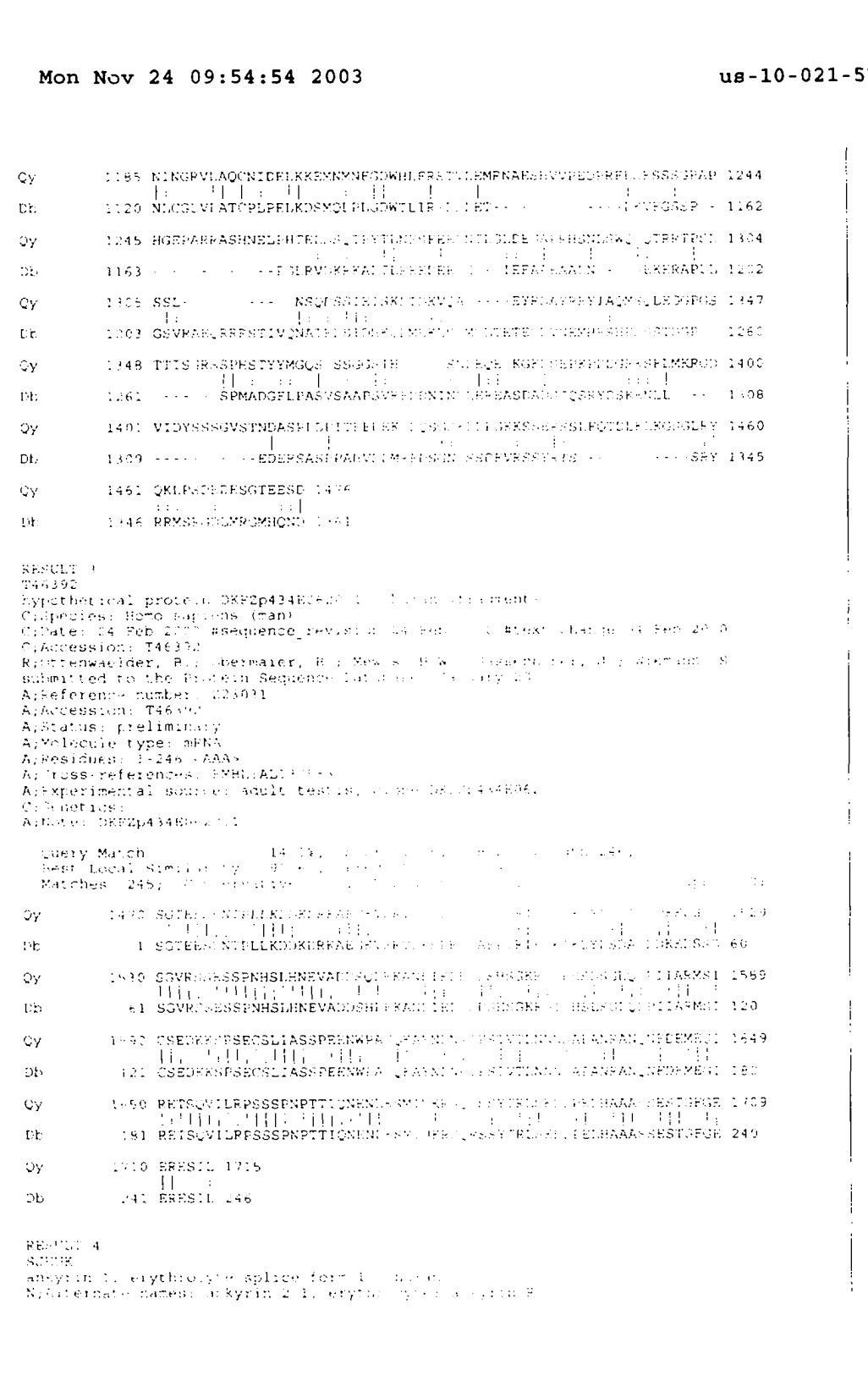
Qy 361 KVSAVDEKGDTPLHIAIRGRSRPKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
Db 361 KVSAVDEKGDTPLHIAIRGRSRPKLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Qy 421 EQIEGAAHLSIETDGHWLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKL 480



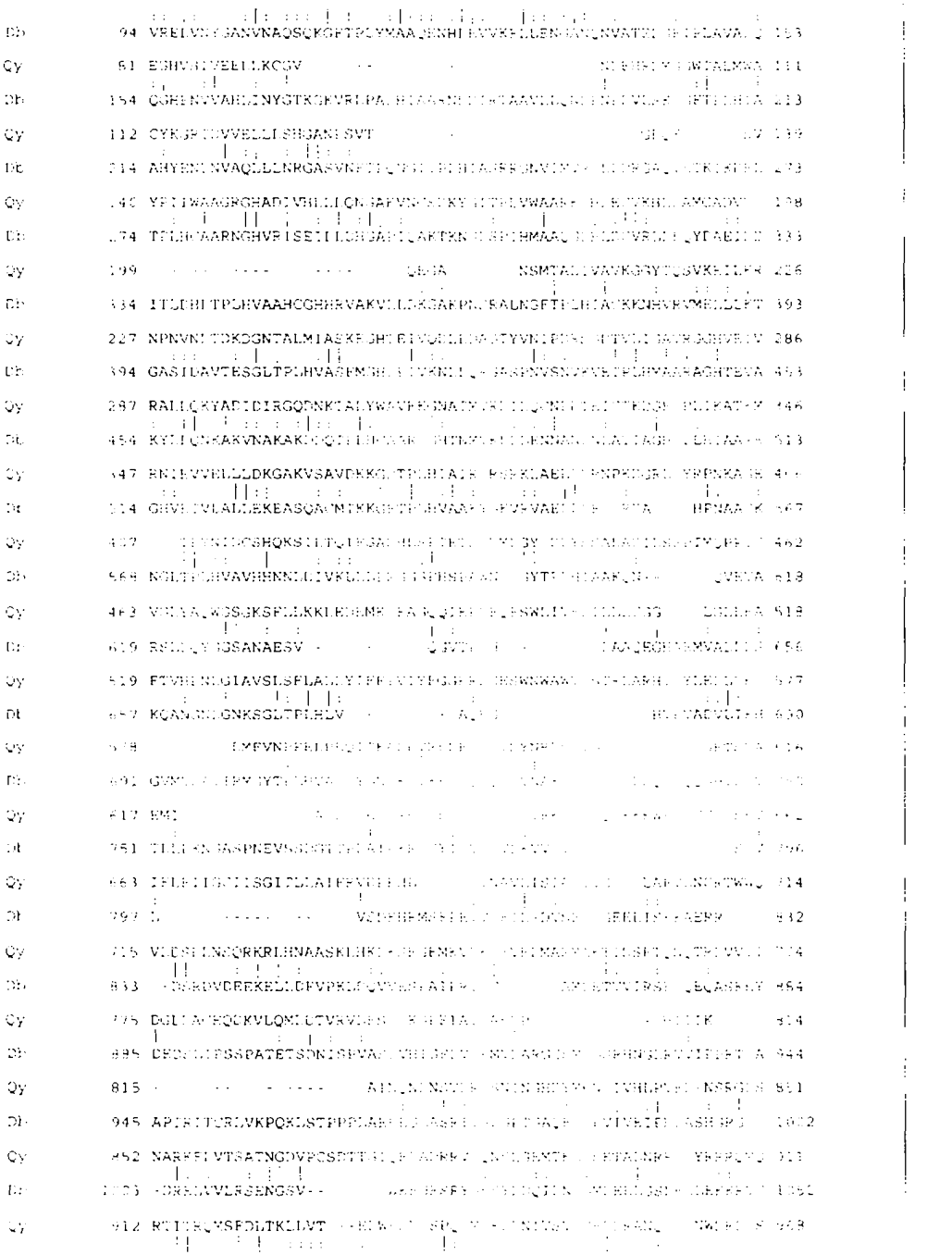
RESULT 2
T21884
hypothetical pr...
C;Species: Cam...
C;Date: 19-Oct-...
C;Accession: T21884
R;Kershaw, J.
submitted to the EMBL Data Library, ...
A;Reference number: D29482
A;Accession: T21884
A;Status: preliminary; translated fr...
A;Molecule type: DNA
A;Residues: 1-1398 <WIL>
A;Cross-references: EMBL:Z68...
A;Experimental sou...: clone P16H...
C;Genetics
A;Gene: CESP:F36H1...
A;Map position: 4
A;Introns: 38/3; ...

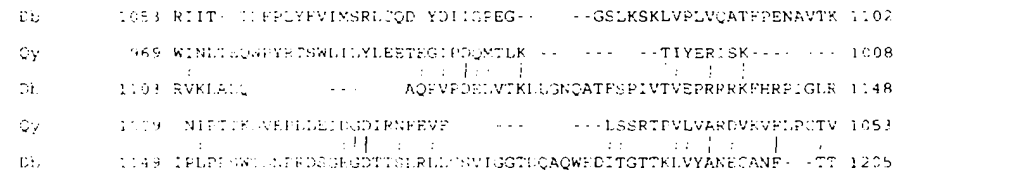
Query Match 20.6%; Score ...
Post Local Similarity 30.9%; Pr...
Matches 468; Conservative 28...



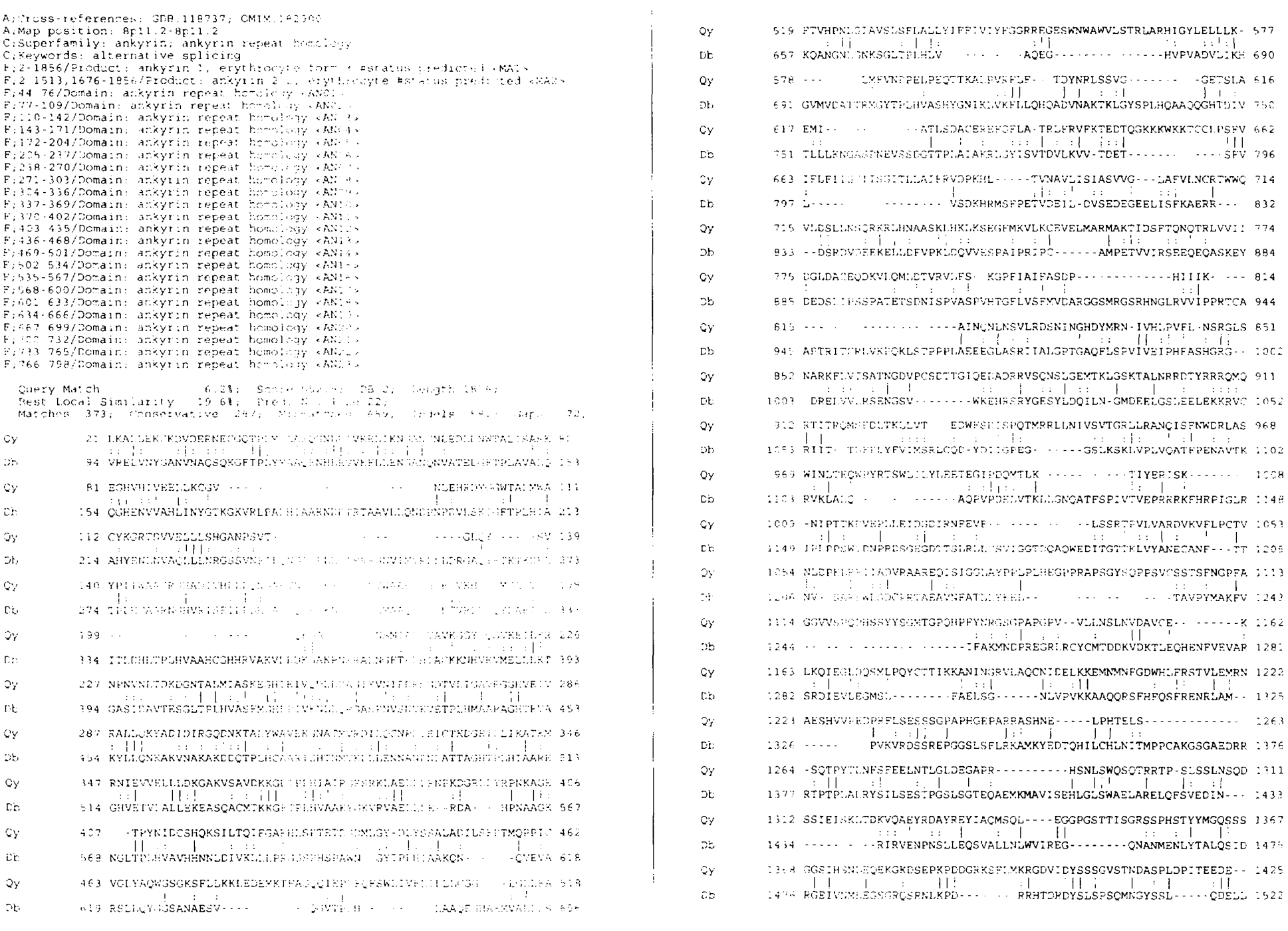
N;Contains: ankyrin 2.2
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22 Jun 1999
C;Accession: S08275; A33219; PC2220; A35443
R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A;Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure
A;Reference number: S08275; MUID:90158830; PMID:2137557
A;Accession: S08275
A;Molecule type: mRNA
A;Residues: 1-1881 <LUX>
A;Cross-references: EMBL:X15629; NID:g28701; PIDN:CAA34610.1; PID:g28702
A;Accession: A33219
A;Molecule type: protein
A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,'X',1367;1383-1427;1431-1630;1696-1699,'D',1700;1763-1772 <LUX>
A;Note: 846-Arg and 1-2 Thr were also found
R;Hermann, J.; Barel, Y.; Prado, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membra
A;Reference number: PC2220; MUID:95071348; PMID:7526850
A;Accession: PC2220
A;Molecule type: protein
A;Residues: 910-929 <HER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchange
A;Reference number: A35443; MUID:90286190; PMID:2141335
A;Accession: A35443
A;Molecule type: protein
A;Residues: 'X',5,'X',7-12;403-412,'X',419-422,'H',424,'LQ';797-800,'E',802-814,862-86
C;Genetics:
A;Gene: GDB:ANK1; ANP
A;Cross-references: GDB:118773; OMIM:182900
A;Map position: 8p11.1-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; phosphoprotein
F;2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F;2-1512,1635-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>
F;2-827/Domain: 89K #status predicted <DOM1>
F;2-827/Region: anion exchange protein binding
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;291-323/Domain: ankyrin repeat homology <AN08>
F;324-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-469/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>
F;828-1382/Domain: NCP #status predicted <DOM2>
F;828-1382/Region: spectrin binding
F;1383-1881/Domain: CBR #status predicted <DOM3>

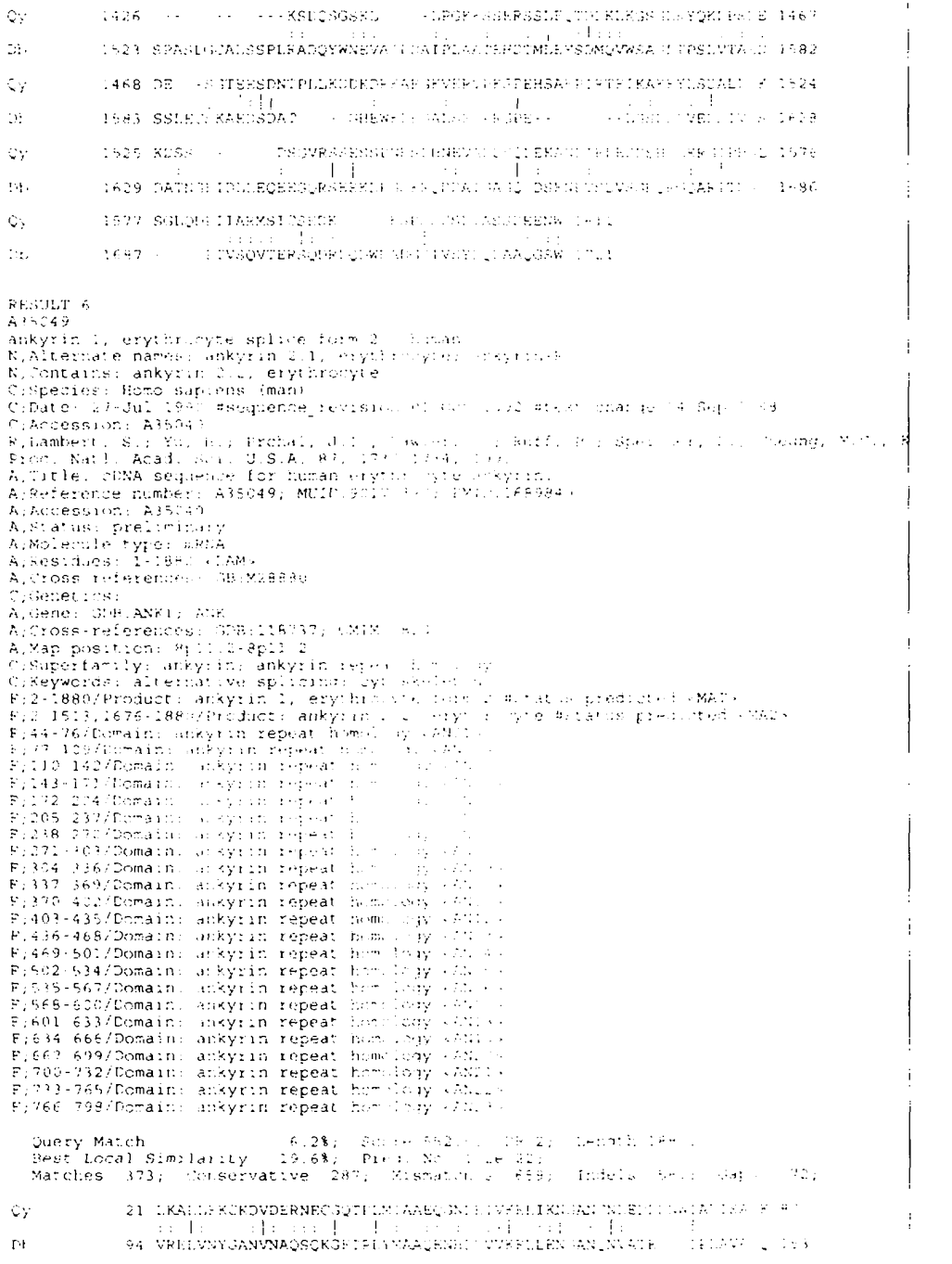
Query Match 6.3%; Score 355.5; DB 1; Length 1881;
Best Local Similarity 20.7%; Pred. No. 8.2e-23;
Matches 424; Conservative 286; Mismatches 669; Indels 581; Gaps 72;





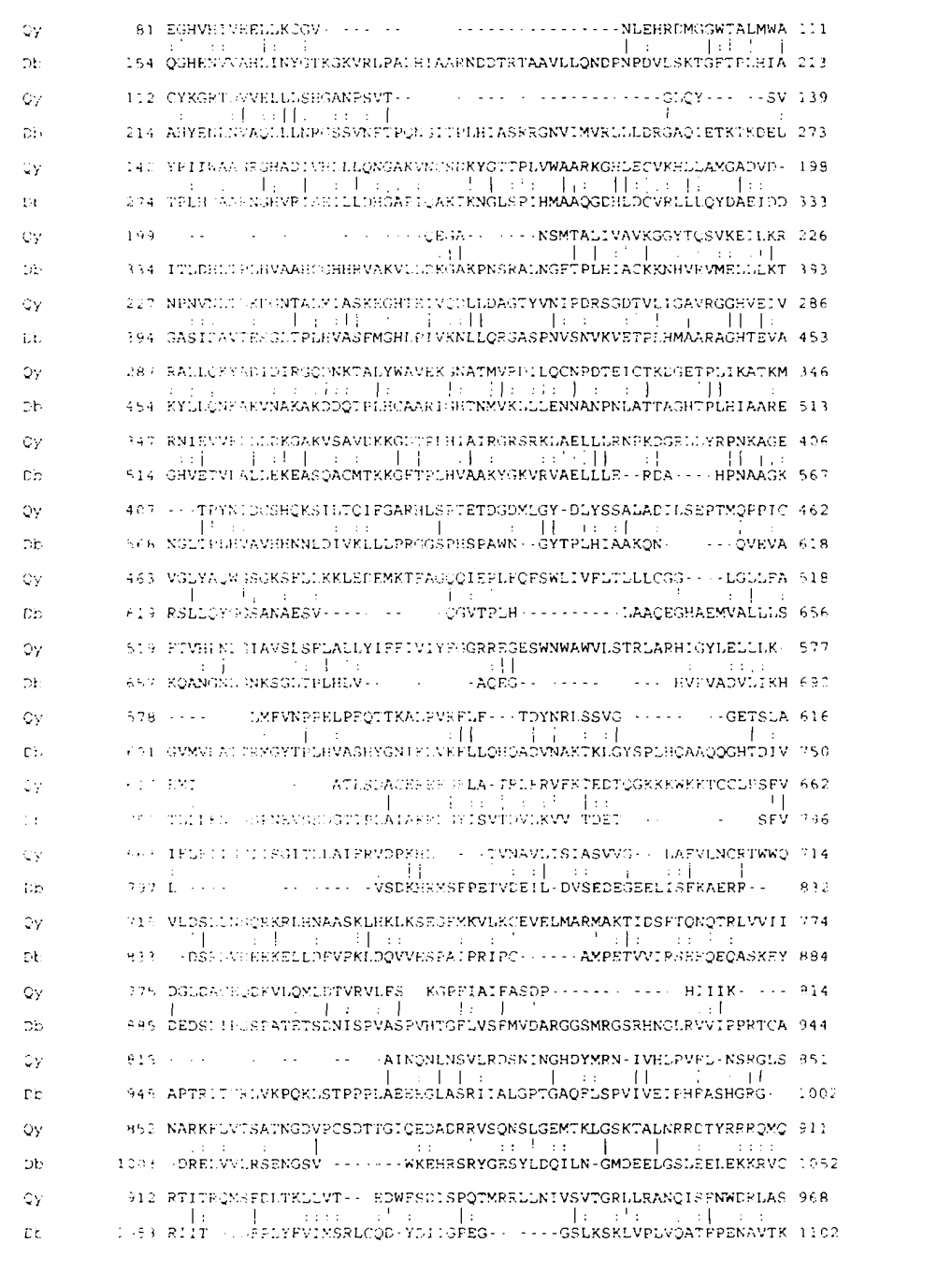
RESULT 5
B35049
ankyrin 1, erythrocyte splice form 3 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 1.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C;Accession: B35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35043; MUID:90175370; PMID:1689849
A;Accession: B35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1896 <LAM>
C;Comments:
A;Gene: GDB:ANK1; OMI

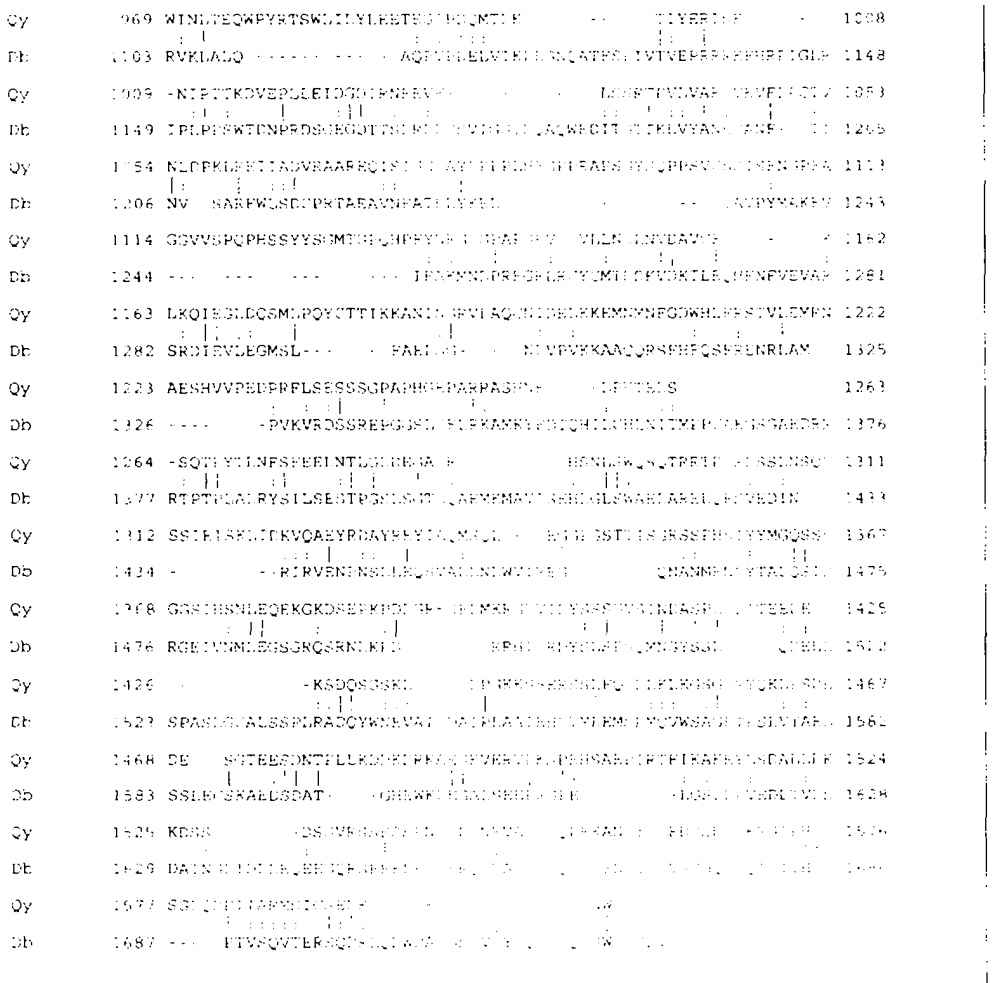




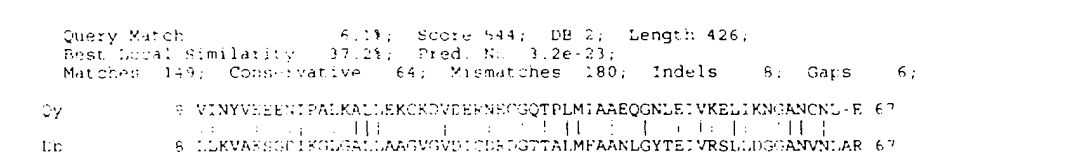
RESULT 6
A35049
ankyrin 1, erythrocyte splice form 2 - human
N;Alternate names: ankyrin 2.1; erythrocyte ankyrin-b
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 27-Jul 199[illegible] #sequence_revision [illegible] #text_change [illegible] Suppl 98
C;Accession: A35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; [illegible]; Kuff, E.; Speicher, D.; Cheung, M.C.; Marchesi, V.T.; Morrow, J.S.
Proc. Natl. Acad. Sci. U.S.A. 88, 1771-1775, 1991
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:91181044; PMID:1672044
A;Accession: A35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1884 <LAM>
A;Cross-references: GB:M28880
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:116737; OMIM:[illegible]
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat family
C;Keywords: alternative splicing; cytoskeleton
F;2-1880/Product: ankyrin 2, erythrocyte form 2 #status predicted <MAT>
F;2-1513,1676-1884/Product: ankyrin 2.2, erythrocyte #status predicted <MAT>
F;44-76/Domain: ankyrin repeat homology <ANK 1>
F;77-109/Domain: ankyrin repeat homology <ANK 2>
F;110-142/Domain: ankyrin repeat homology <ANK 3>
F;143-172/Domain: ankyrin repeat homology <ANK 4>
F;172-204/Domain: ankyrin repeat homology <ANK 5>
F;205-237/Domain: ankyrin repeat homology <ANK 6>
F;238-270/Domain: ankyrin repeat homology <ANK 7>
F;271-303/Domain: ankyrin repeat homology <ANK 8>
F;304-336/Domain: ankyrin repeat homology <ANK 9>
F;337-369/Domain: ankyrin repeat homology <ANK 10>
F;370-402/Domain: ankyrin repeat homology <ANK 11>
F;403-435/Domain: ankyrin repeat homology <ANK 12>
F;436-468/Domain: ankyrin repeat homology <ANK 13>
F;469-501/Domain: ankyrin repeat homology <ANK 14>
F;502-534/Domain: ankyrin repeat homology <ANK 15>
F;535-567/Domain: ankyrin repeat homology <ANK 16>
F;568-600/Domain: ankyrin repeat homology <ANK 17>
F;601-633/Domain: ankyrin repeat homology <ANK 18>
F;634-666/Domain: ankyrin repeat homology <ANK 19>
F;667-699/Domain: ankyrin repeat homology <ANK 20>
F;700-732/Domain: ankyrin repeat homology <ANK 21>
F;733-765/Domain: ankyrin repeat homology <ANK 22>
F;766-799/Domain: ankyrin repeat homology <ANK 23>

Query Match 6.2%; Score 562.[illegible]; DB 2; Length 1884;
Best Local Similarity 19.6%; Pred. No. [illegible]e-81;
Matches 373; Conservative 280; Mismatches 859; Indels [illegible]; Gaps [illegible];

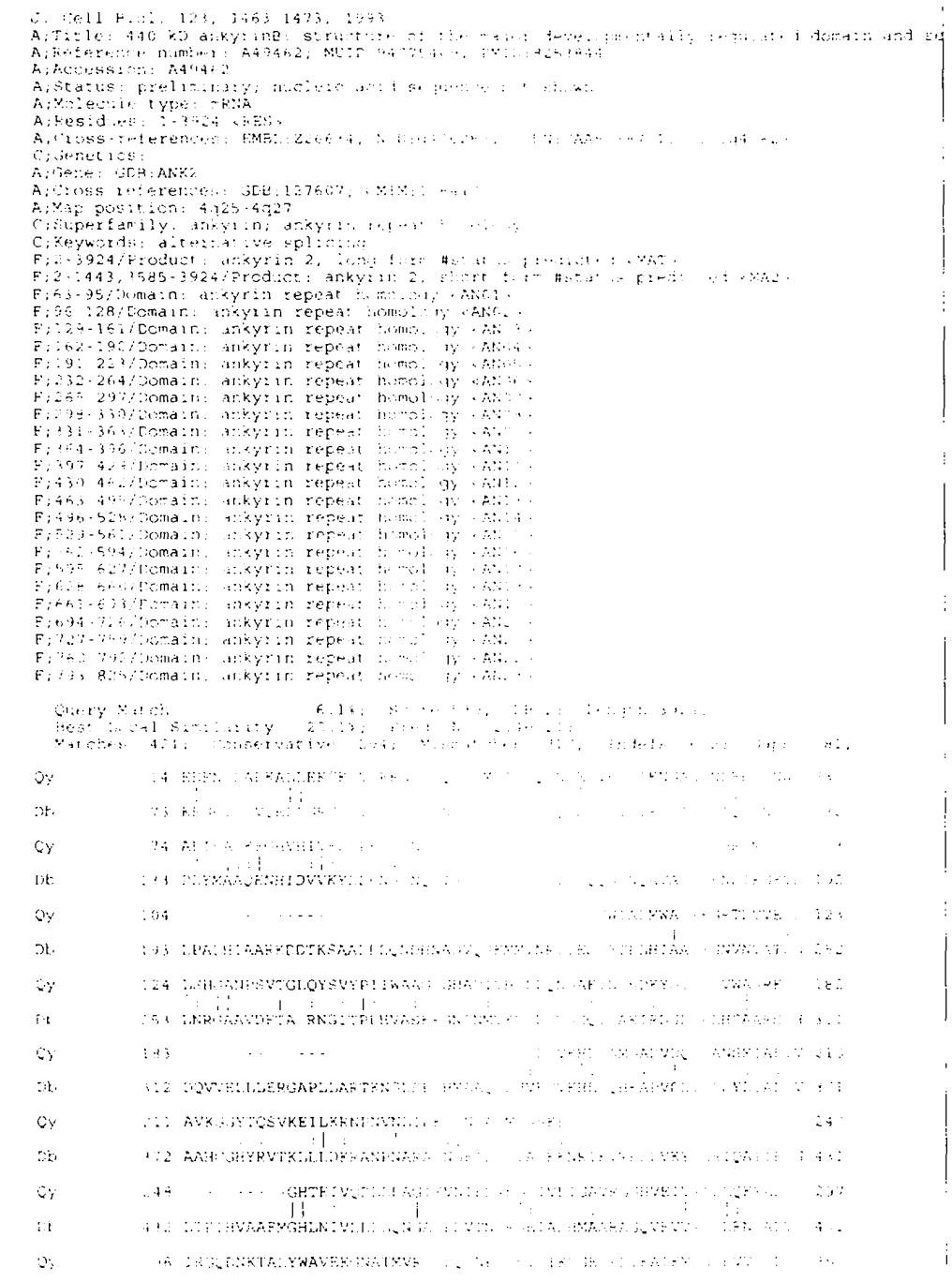


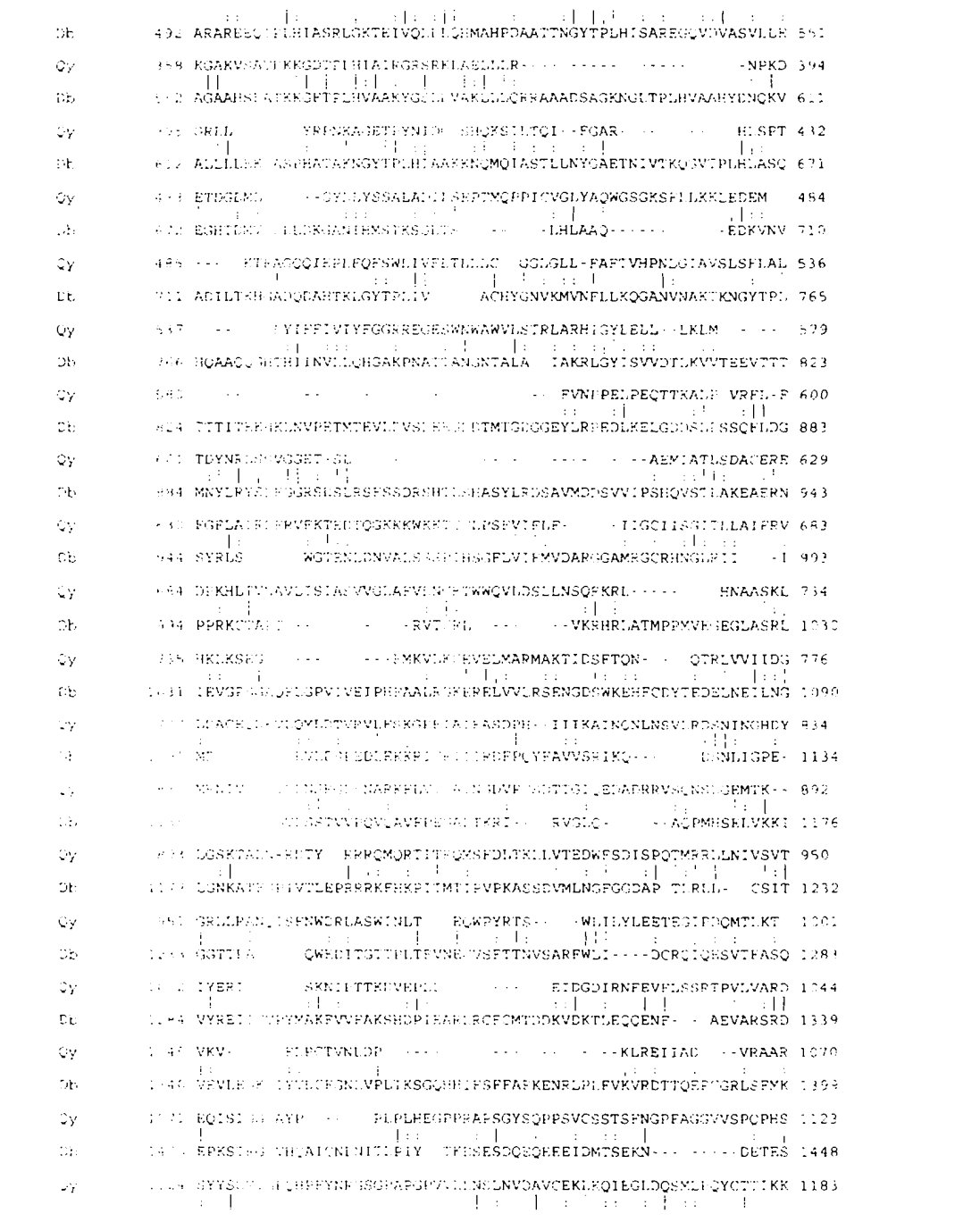


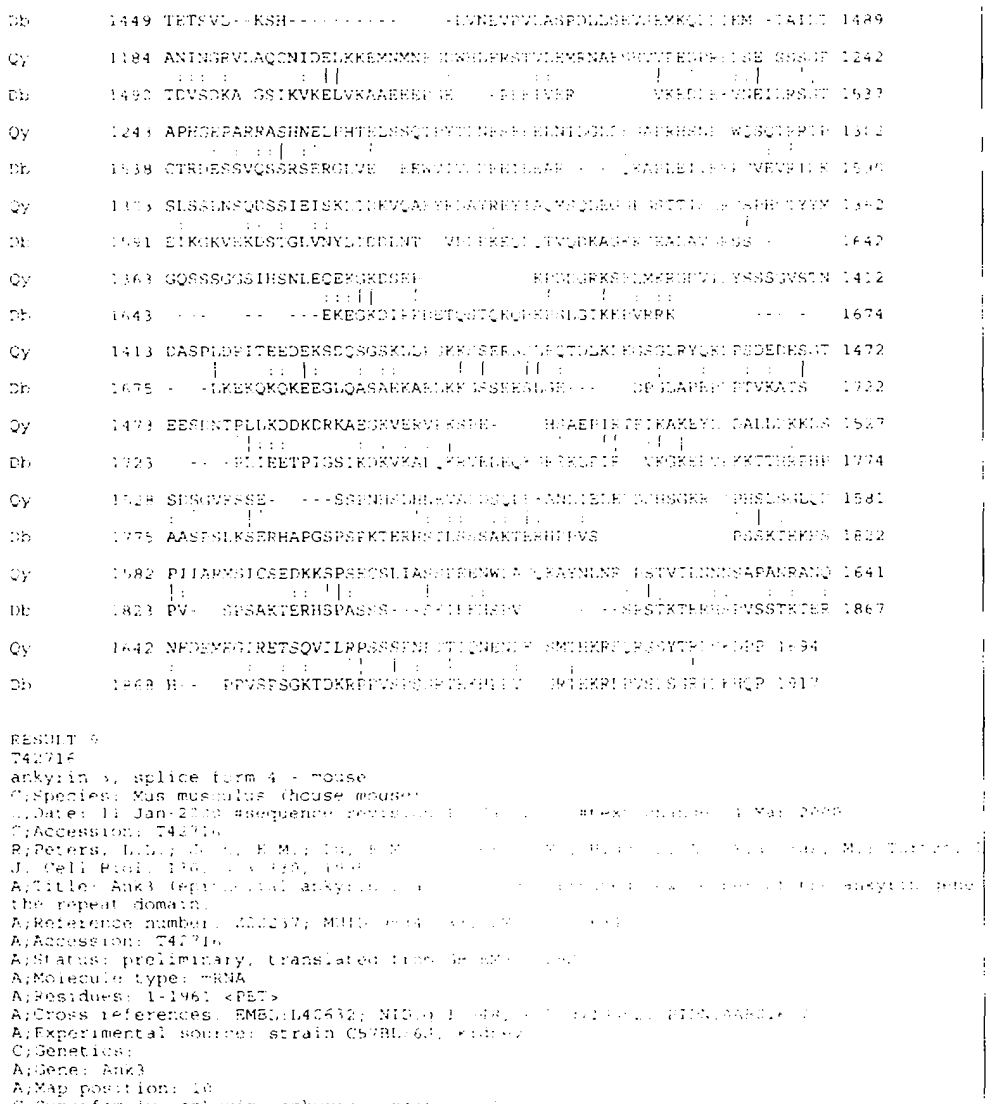
RESULT 7
AE2149
hypothetical protein all2748 [imported - Nostoc sp. PCC 7120
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 [imported - Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 27-Jul-2002
C;Accession: AE2149
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasaki, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Hirata, M.; Yamada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AR1807; MUID:21[illegible]; PMID:[illegible]
A;Accession: AE2149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-426 <KUR>
A;Cross-references: GB:BA000019; PID:[illegible]
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2748



RESULT 8
S37431
ankyrin 2, neuronal long splice form - human
N;Alternate names: ankyrin B, 440K splice form; ankyrin B; brain ankyrin; non-erythroi
N;Contains: ankyrin 2, short form
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug 1999
C;Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
R;Chan, W.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1924 <CHA>
A;Cross-references: EMBL:Z26634; NID:g407147; PIDN:CAA81387.1; PID:g406288
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal
A;Reference number: A39643; MUID:91302466; PMID:1830053
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2077 <OTT>
A;Cross-references: EMBL:X56958
A;Accession: B39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443,<1445-1924 <OTT>
A;Cross-references: EMBL:X56958
R;Tse, W.T.; Menninger, J.C.; Yang Feng, T.L.; Francke, U., Sahr, K.E.; Lux, S.E.; War
Genomics 10, 859-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92008921; PMID:1933308
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 47-474,<476-477,<479-496 <TSE>
A;Cross-references: [illegible]; NID:g178447; PIDN:AAA62848.1; PID:g178648
R;Chan, W.; Kunimoto, M.; Otto, E.

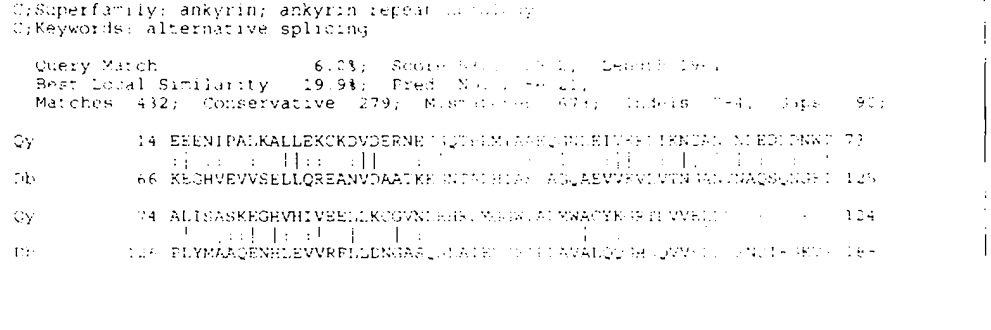


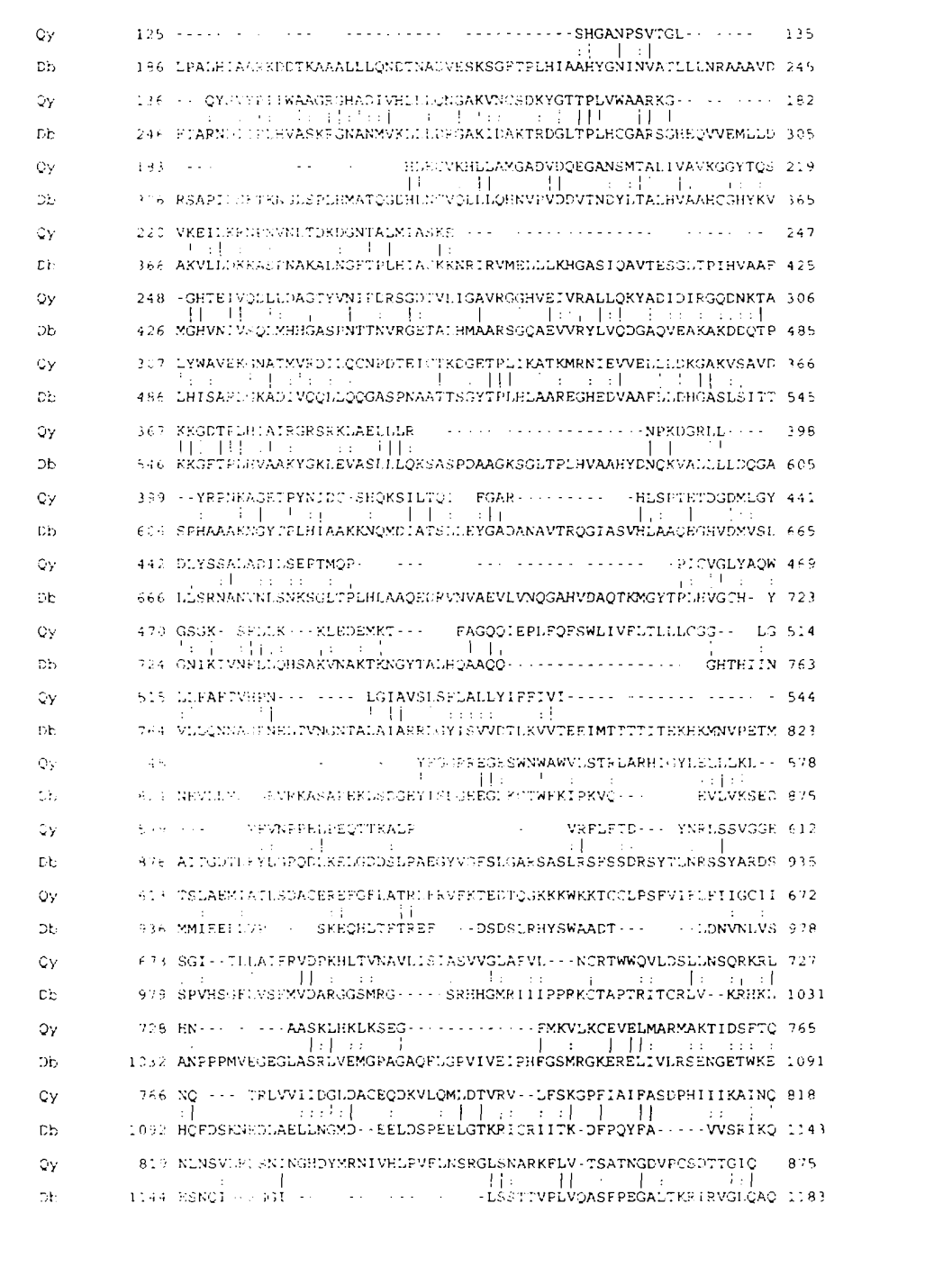


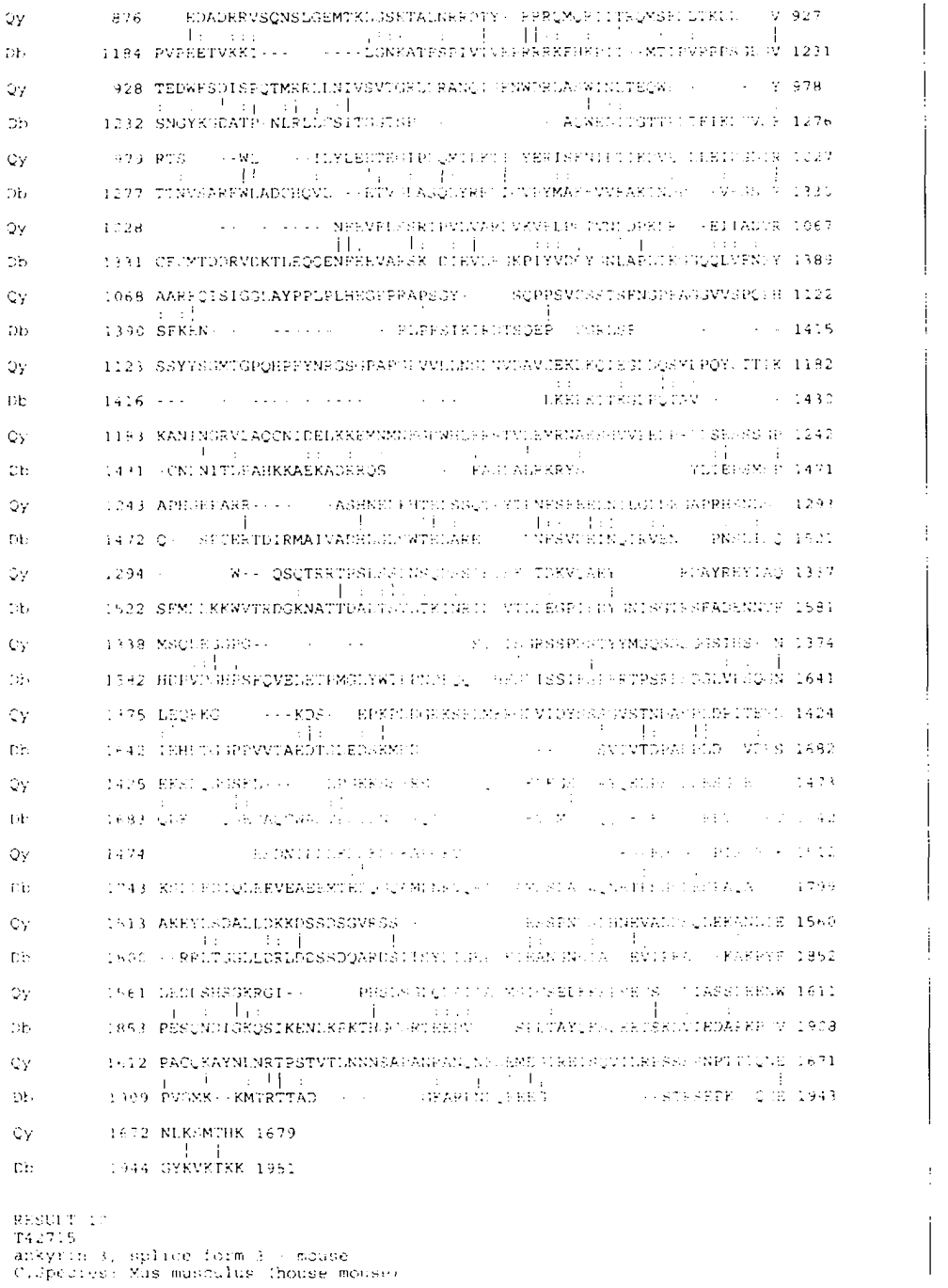


RESULT 9
T42716
ankyrin 3, splice form 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence revision [illegible] 1-Mar-2000
C;Accession: T42716
R;Peters, L.L.; [illegible] E.M.; Inc. B.W. [illegible] Nature Genet. M.; [illegible]
J. Cell Biol. 136, [illegible], 1997
A;Title: Ank3 (epithelial ankyrin [illegible] ankyrin gene
the repeat domain.
A;Reference number: A02697; MUID: [illegible]
A;Accession: T42716
A;Status: preliminary; translated from GP
A;Molecule type: mRNA
A;Residues: 1-1961 <PET>
A;Cross-references: EMBL:L40632; NID:g [illegible]
A;Experimental source: strain C57BL/6J, [illegible]
C;Genetics:
A;Gene: Ank3
A;Map position: 10
C;Superfamily: ankyrin; ankyrin repeat [illegible]
C;Keywords: alternative splicing

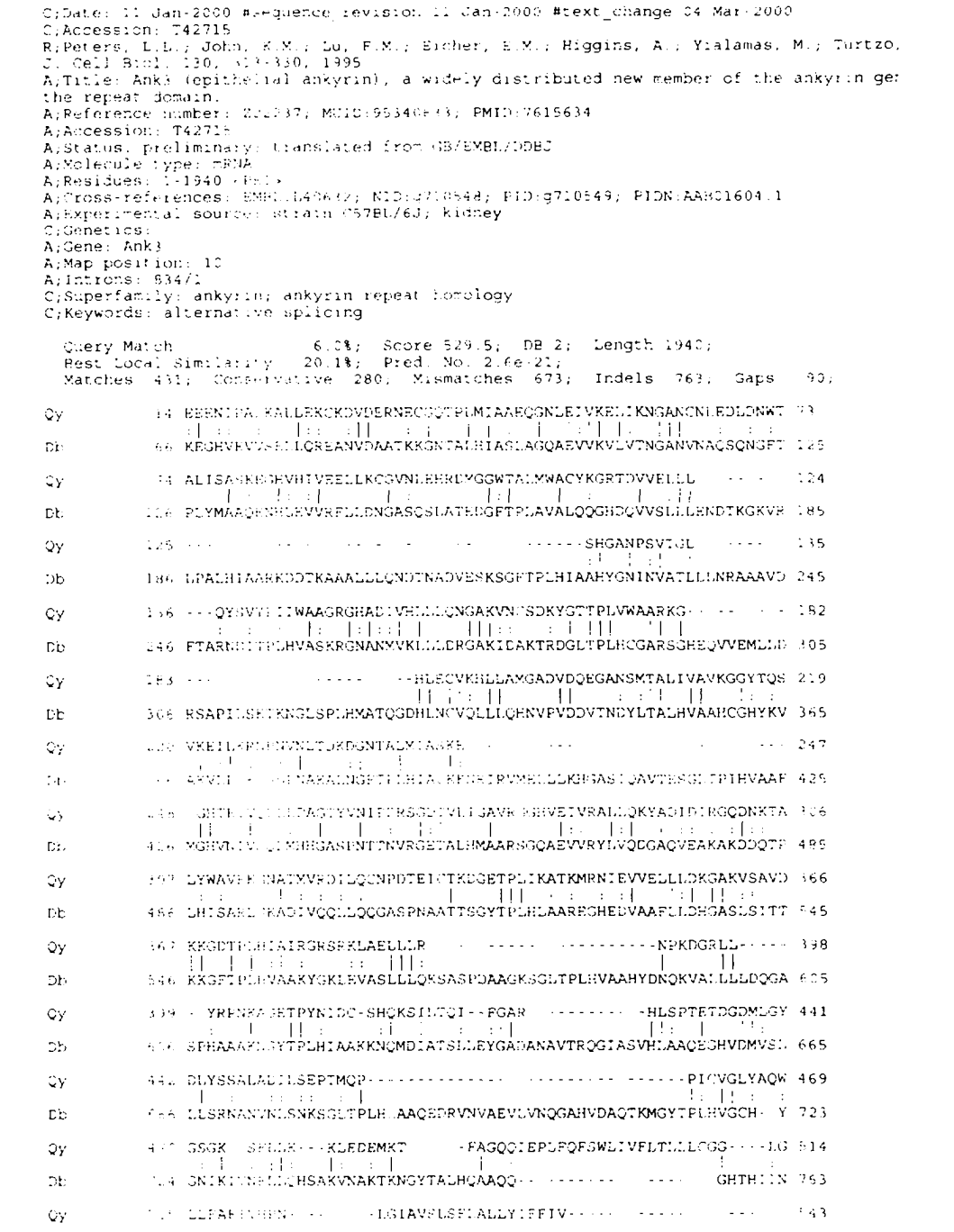
Query Match 6.0%; Score [illegible]
Best Local Similarity 19.9%; Pred [illegible]
Matches 432; Conservative 279; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

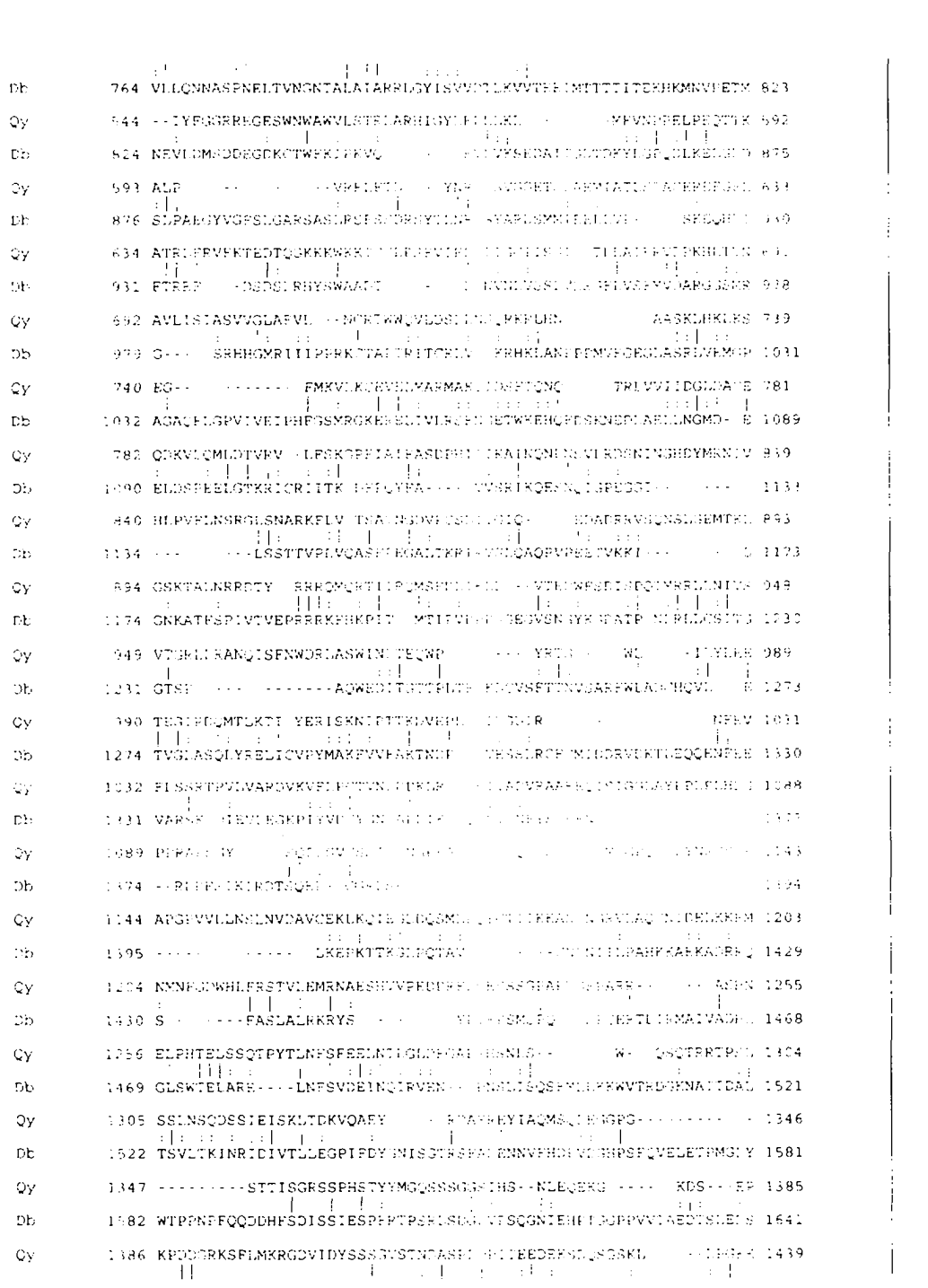






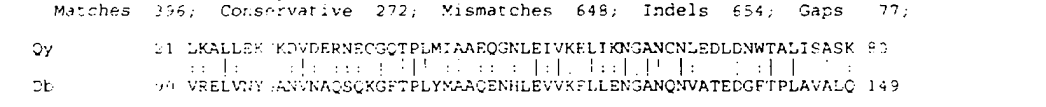
RESULT 17
T42715
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, C. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge the repeat domain.
A;Reference number: A42387; MUID:95340643; PMID:7615634
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <Pet>
A;Cross-references: EMBL:L40642; NID:g710548; PID:g710549; PIDN:AAA01604.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 834/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

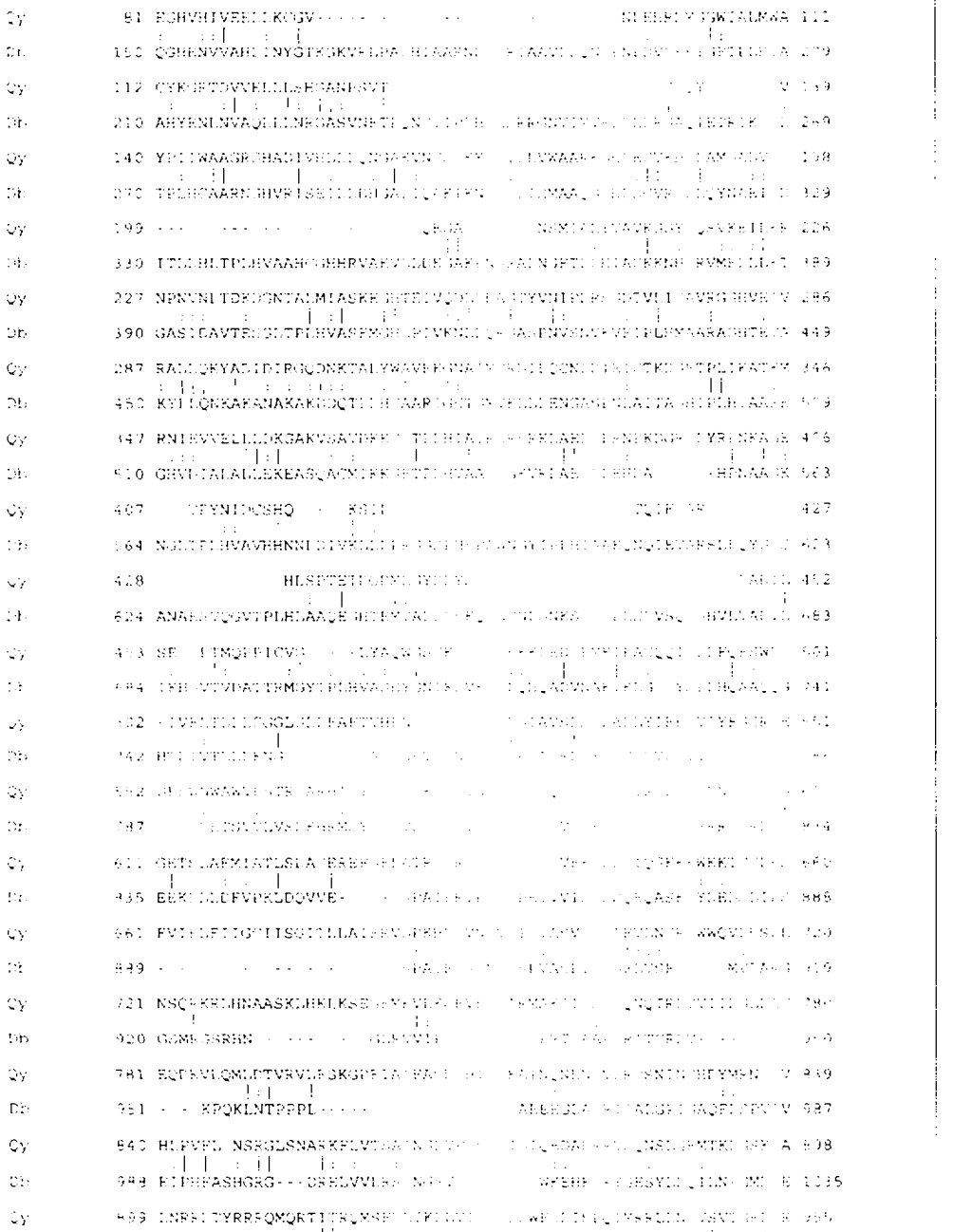


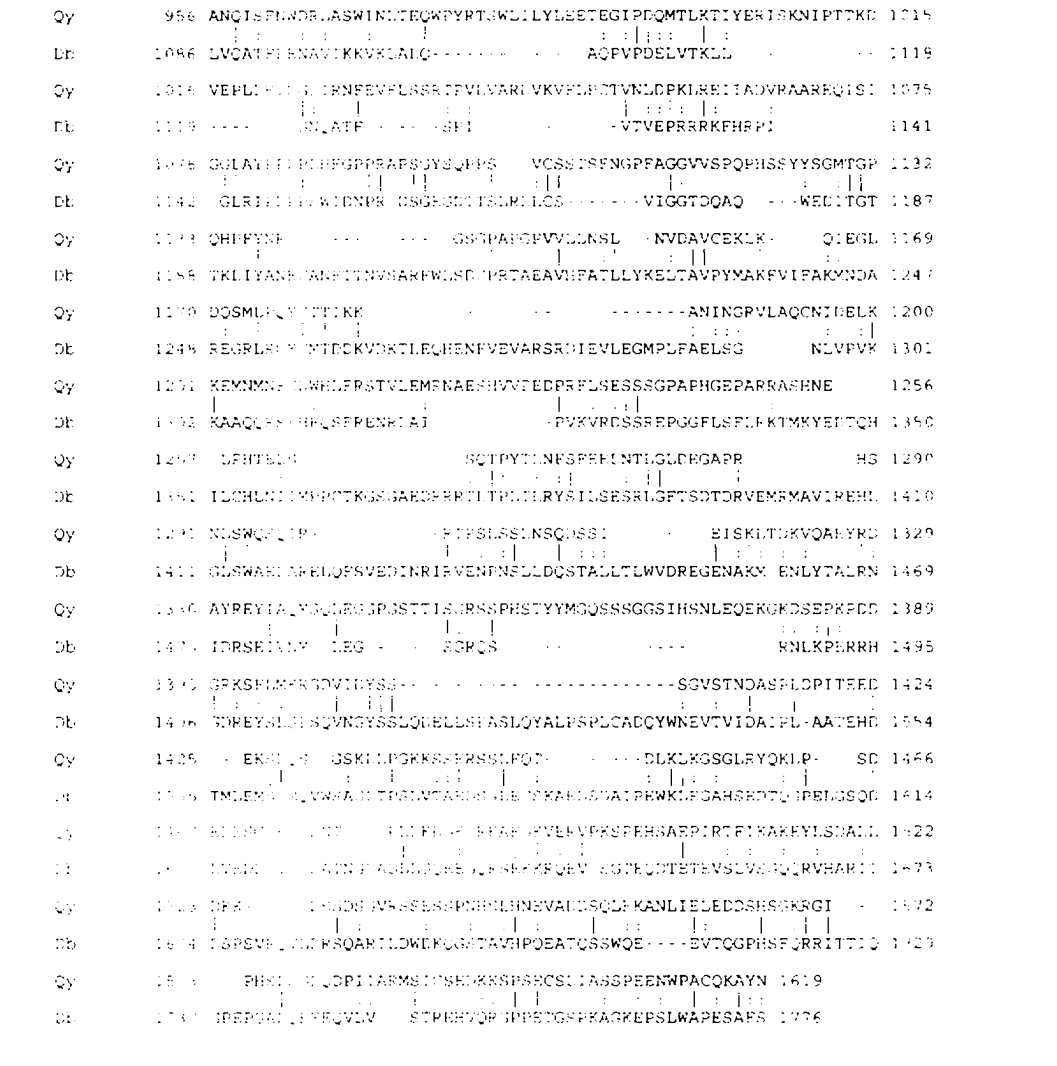


RESULT 11
I49502
ankyrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C;Accession: I49502
R;White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A;Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory
A;Reference number: I49502; MUID:92345717; PMID:1386265
A;Accession: I49502
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1862 <PES>
A;Cross-references: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C;Genetics:
A;Gene: Ank 1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;40-72/Domain: ankyrin repeat homology <AN01>
F;73-105/Domain: ankyrin repeat homology <AN02>
F;106-138/Domain: ankyrin repeat homology <AN03>
F;139-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;267-299/Domain: ankyrin repeat homology <AN08>
F;300-332/Domain: ankyrin repeat homology <AN09>
F;333-365/Domain: ankyrin repeat homology <AN10>
F;366-398/Domain: ankyrin repeat homology <AN11>
F;399-431/Domain: ankyrin repeat homology <AN12>
F;432-464/Domain: ankyrin repeat homology <AN13>
F;465-497/Domain: ankyrin repeat homology <AN14>
F;498-530/Domain: ankyrin repeat homology <AN15>
F;531-563/Domain: ankyrin repeat homology <AN16>
F;564-596/Domain: ankyrin repeat homology <AN17>
F;597-629/Domain: ankyrin repeat homology <AN18>
F;630-662/Domain: ankyrin repeat homology <AN19>
F;663-695/Domain: ankyrin repeat homology <AN20>
F;696-728/Domain: ankyrin repeat homology <AN21>
F;729-761/Domain: ankyrin repeat homology <AN22>
F;762-794/Domain: ankyrin repeat homology <AN23>

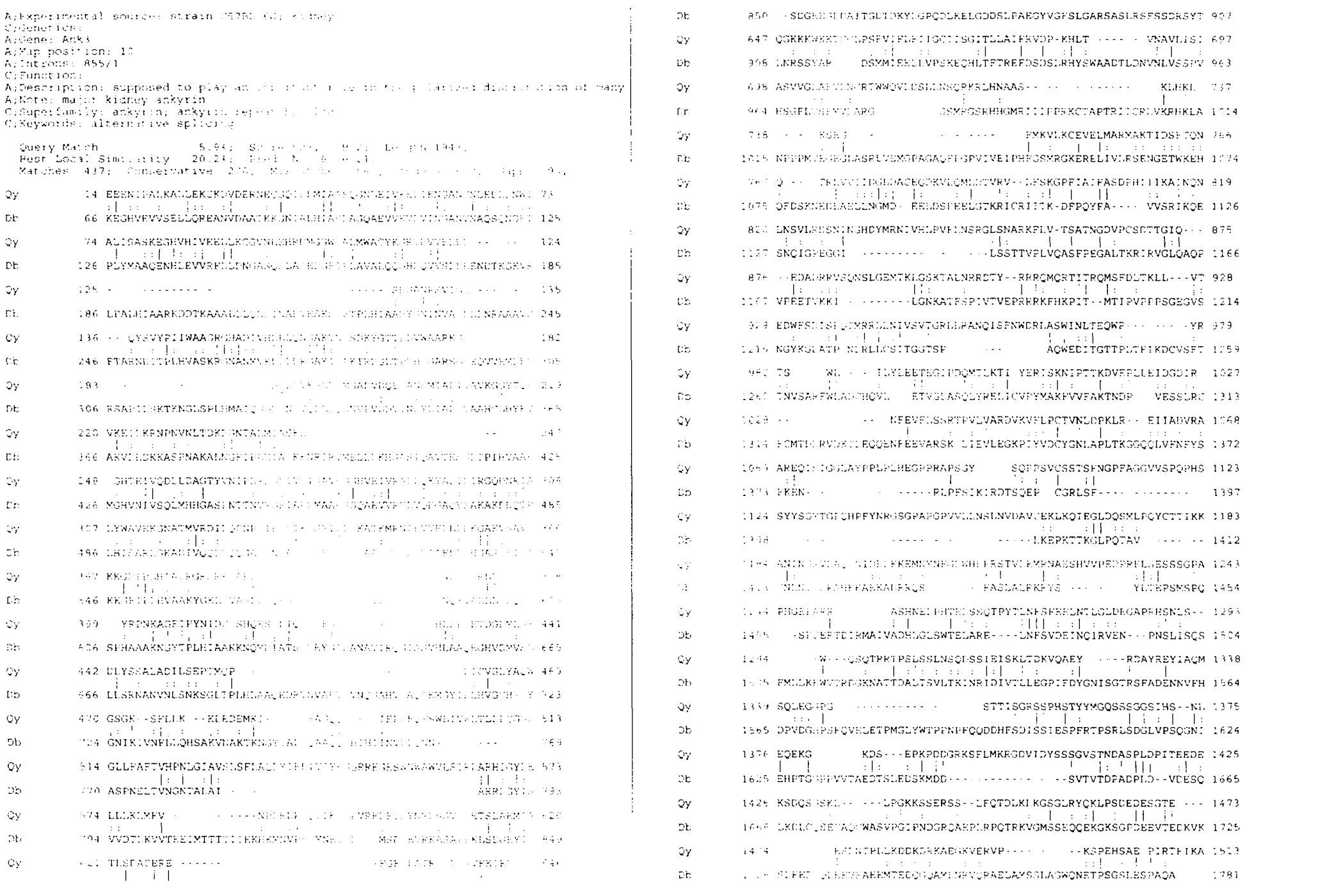
Query Match 5.9%; Score 525; DB 2; Length 1862;
Best Local Similarity 20.1%; Pred. No. 4.4e-21;
Matches 356; Conservative 272; Mismatches 648; Indels 654; Gaps 77;

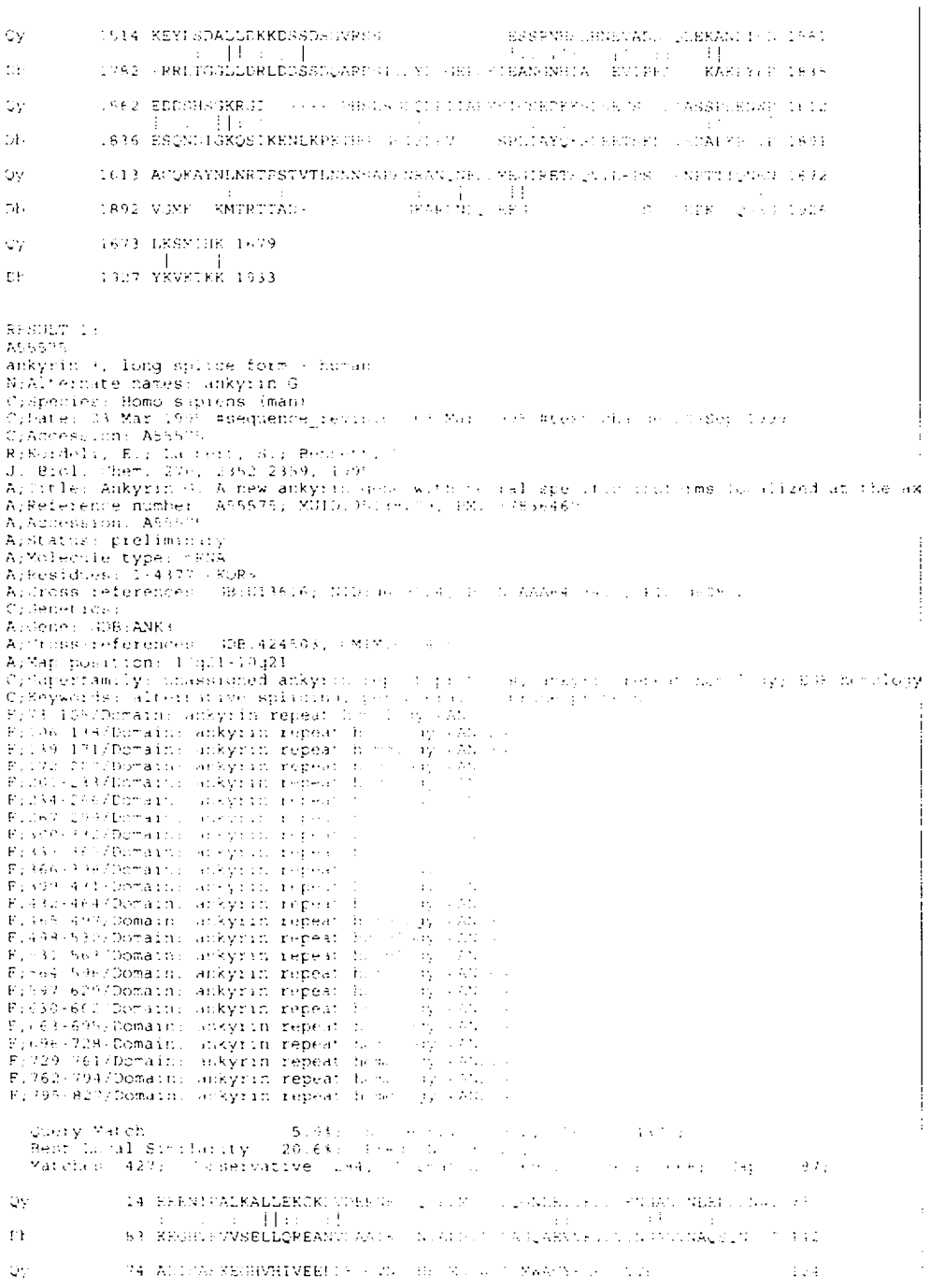






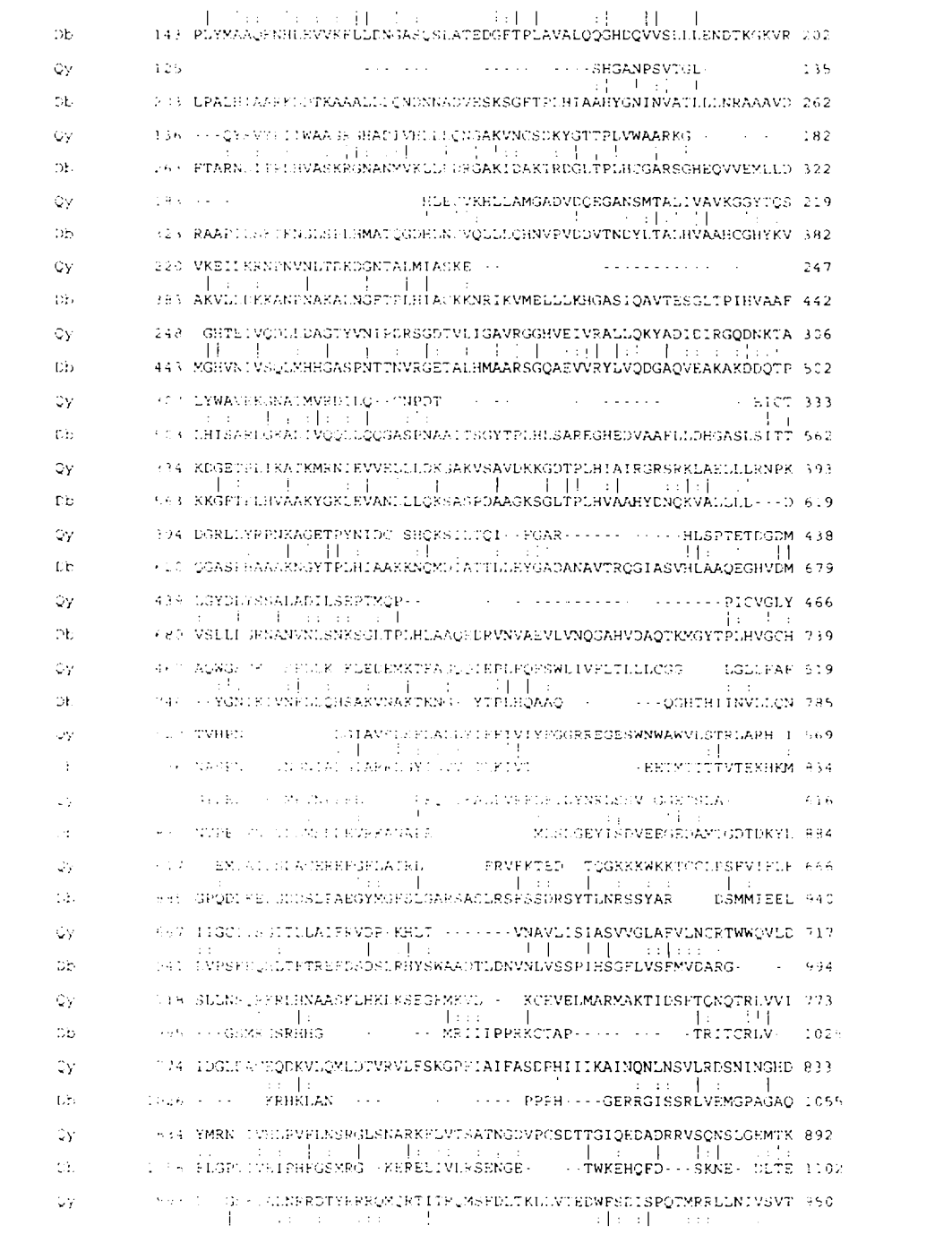
RESULT 11
T42713
ankyrin 3, epithelial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04 Mar-2000
C;Accession: T42713
R;Peters, L.L.; Johnson, K.M.; Lux, S.E.; Hicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo,
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gen
the repeat domain.
A;Reference number: J.2247; MUID:95340643; PMID:7615634
A;Accession: T42713
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1943 <PET>
A;Cross-references: GB:U70649; NID:g710650; PIDN:AAA016<6.1

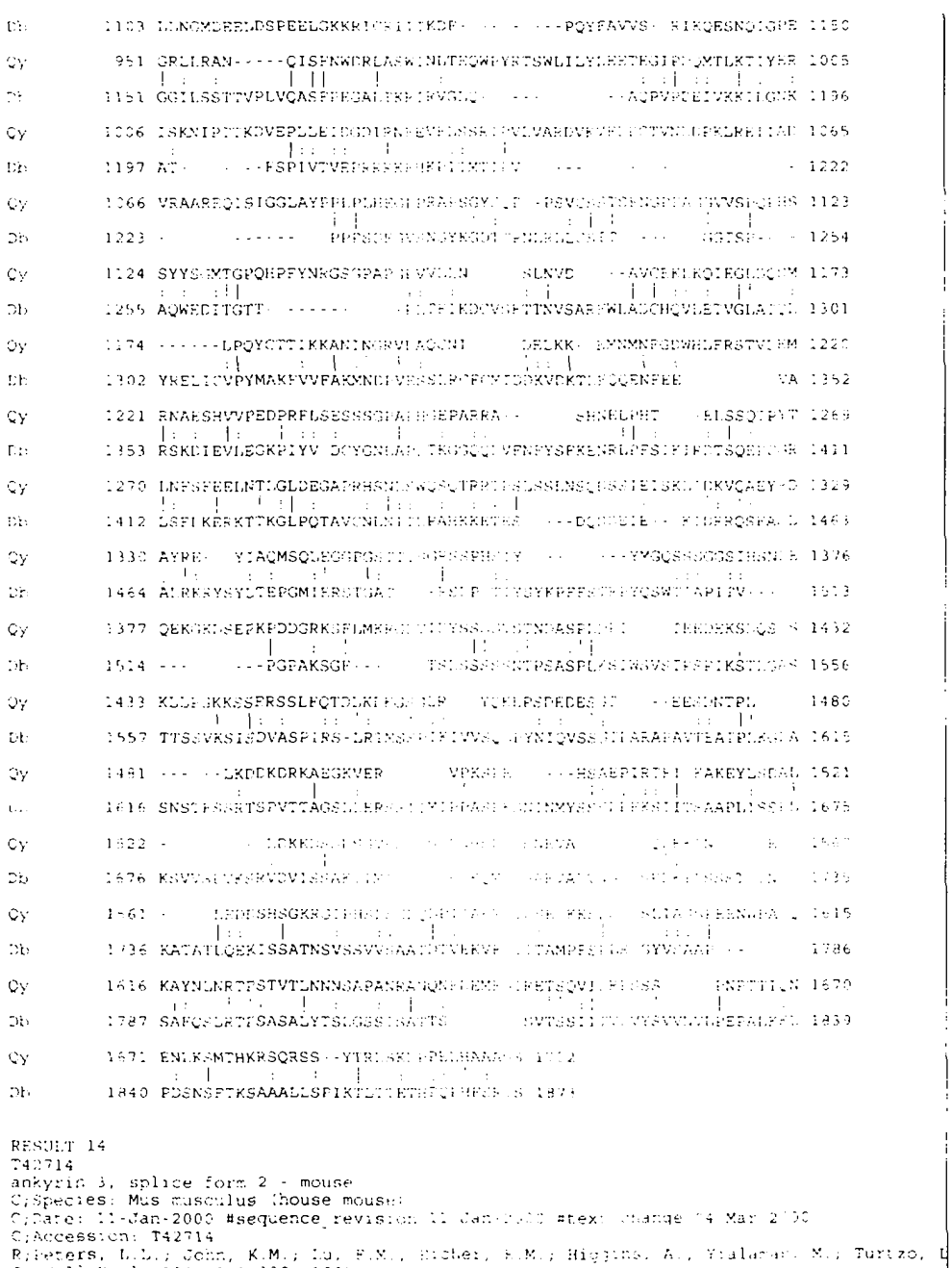




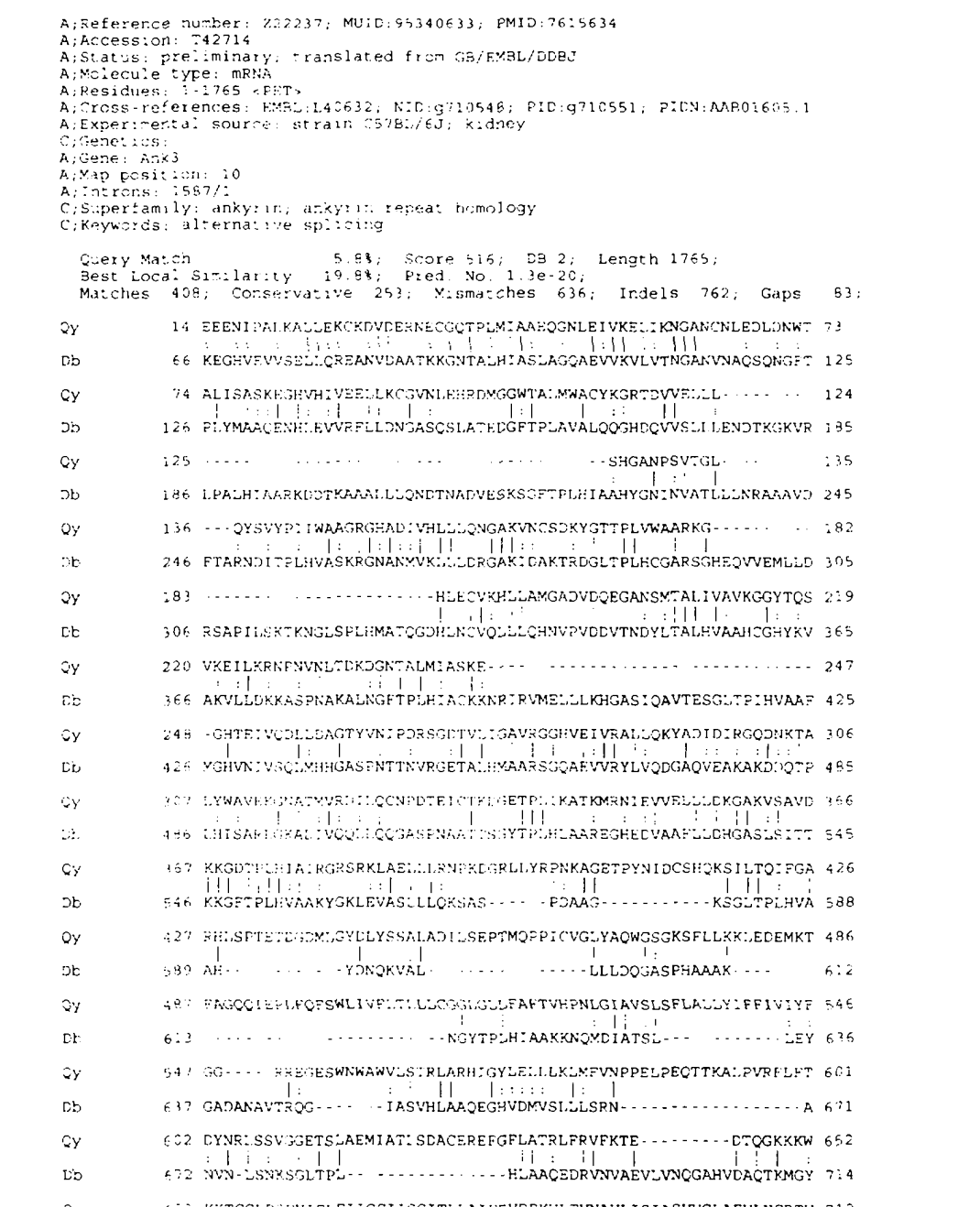
RESULT 1
A55575
ankyrin 4, long splice form - human
N;Alternate names: ankyrin G
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1998 #sequence_revision 23-Mar-1998 #text_change 09-Jul-1999
C;Accession: A55575
R;Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A;Title: Ankyrin G. A new ankyrin gene with neural-specific isoforms localized at the ax
A;Reference number: A55575; MUID:95126446; PMID:7836456
A;Accession: A55575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4377 <KOR>
A;Cross-references: GB:U03136.6; GB:U03137.6; GB:AA044444; NID:g462; PID:g462
C;Genetics:
A;Gene: GDB:ANK4
A;Cross-reference: GDB:424604, OMIM:600
A;Map position: 1p13.1-q21
C;Superfamily: unassigned ankyrin repeat; unassigned ankyrin repeat homology
C;Keywords: alternative splicing; cytoskeleton; membrane; neural
F;23-136/Domain: ankyrin repeat 1
F;139-171/Domain: ankyrin repeat 2
F;172-203/Domain: ankyrin repeat 3
F;206-237/Domain: ankyrin repeat 4
F;254-286/Domain: ankyrin repeat 5
F;287-319/Domain: ankyrin repeat 6
F;320-352/Domain: ankyrin repeat 7
F;353-385/Domain: ankyrin repeat 8
F;386-418/Domain: ankyrin repeat 9
F;419-451/Domain: ankyrin repeat 10
F;452-484/Domain: ankyrin repeat 11
F;485-517/Domain: ankyrin repeat 12
F;488-520/Domain: ankyrin repeat 13
F;521-553/Domain: ankyrin repeat 14
F;554-586/Domain: ankyrin repeat 15
F;587-619/Domain: ankyrin repeat 16
F;620-652/Domain: ankyrin repeat 17
F;653-685/Domain: ankyrin repeat 18
F;686-718/Domain: ankyrin repeat 19
F;729-761/Domain: ankyrin repeat 20
F;762-794/Domain: ankyrin repeat 21
F;795-827/Domain: ankyrin repeat 22

Query Match 5.0%; Score 1
Best Local Similarity 20.6%; Pred. No. 0.
Matches 42%; Conservative 0%; Mismatches 0%; Indels 0%; Gaps 0%



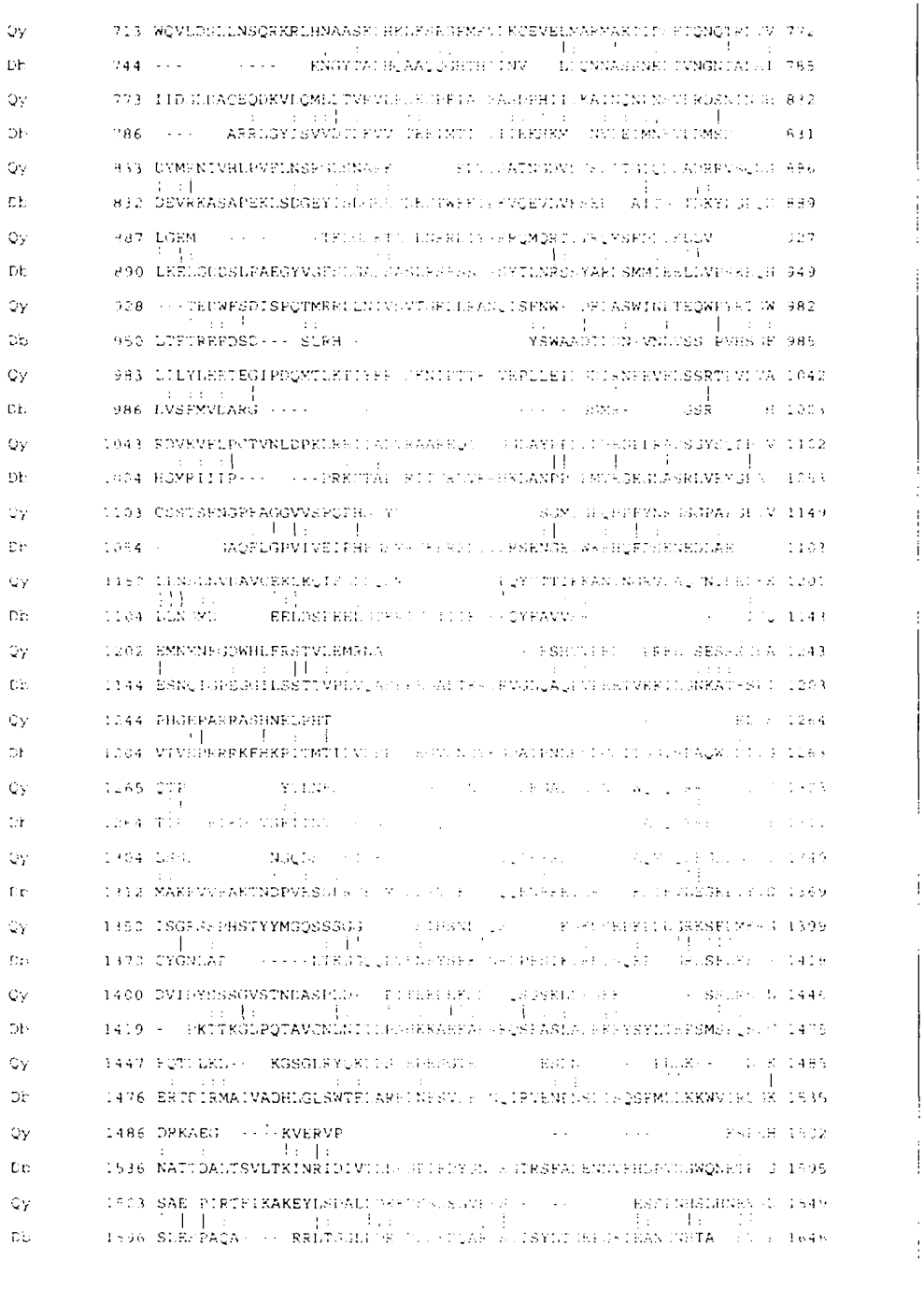


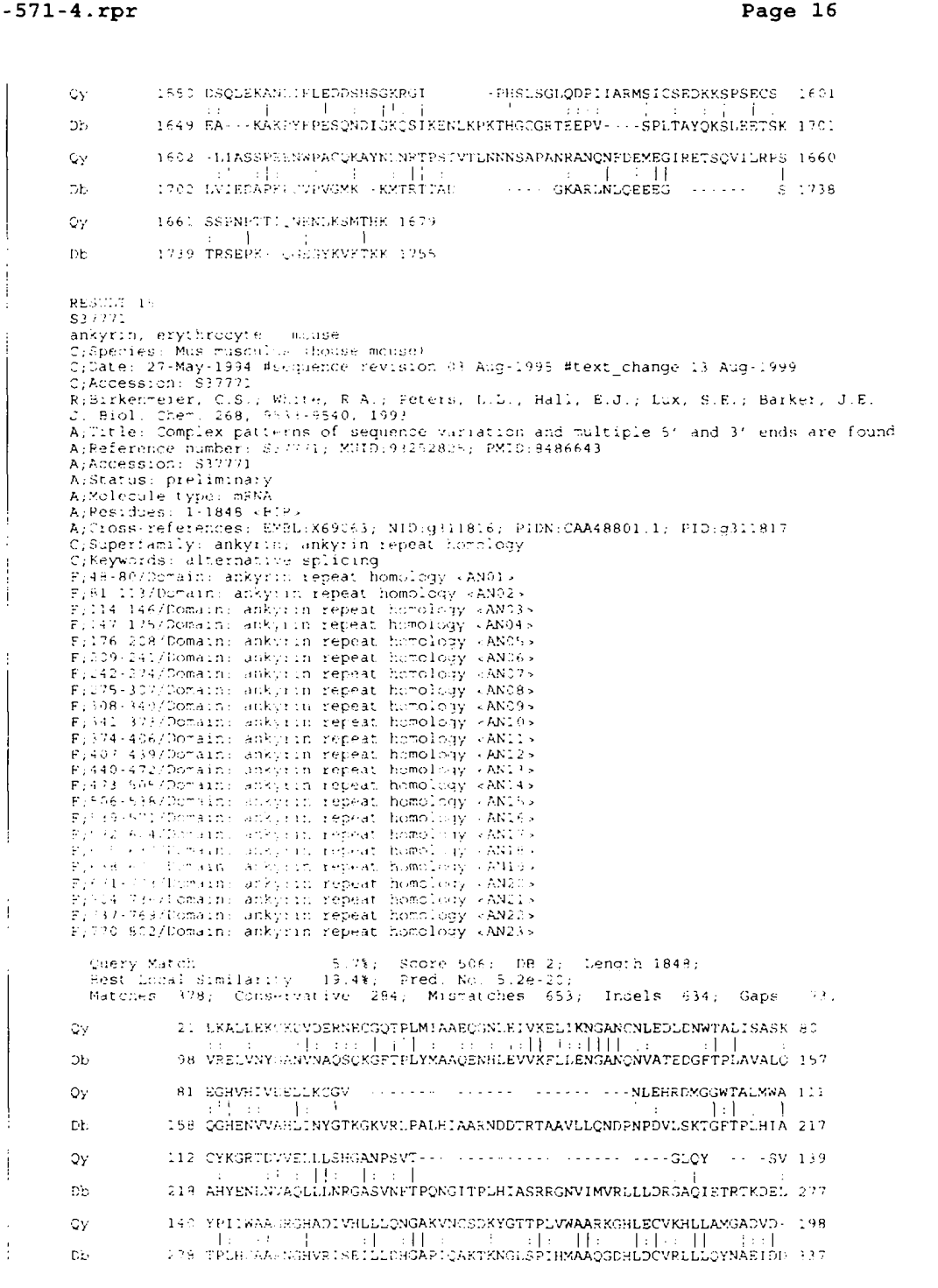
RESULT 14
T42714
ankyrin 3, splice form 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 24 Mar 2000
C;Accession: T42714
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.



A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 1587/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

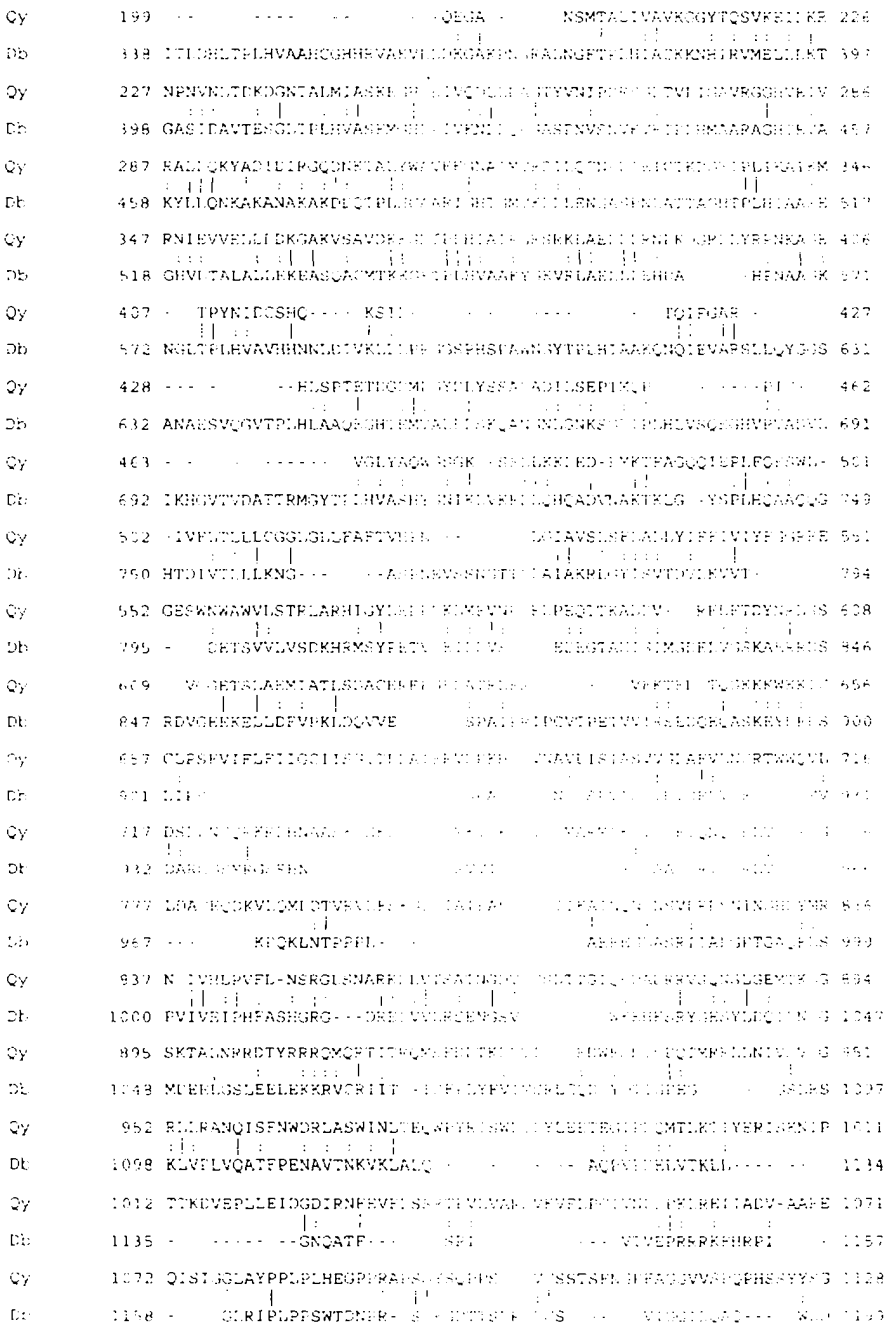
Query Match 5.8%; Score 516; DB 2; Length 1765;
Best Local Similarity 19.8%; Pred. No. 1.3e-20;
Matches 408; Conservative 253; Mismatches 636; Indels 762; Gaps 83;

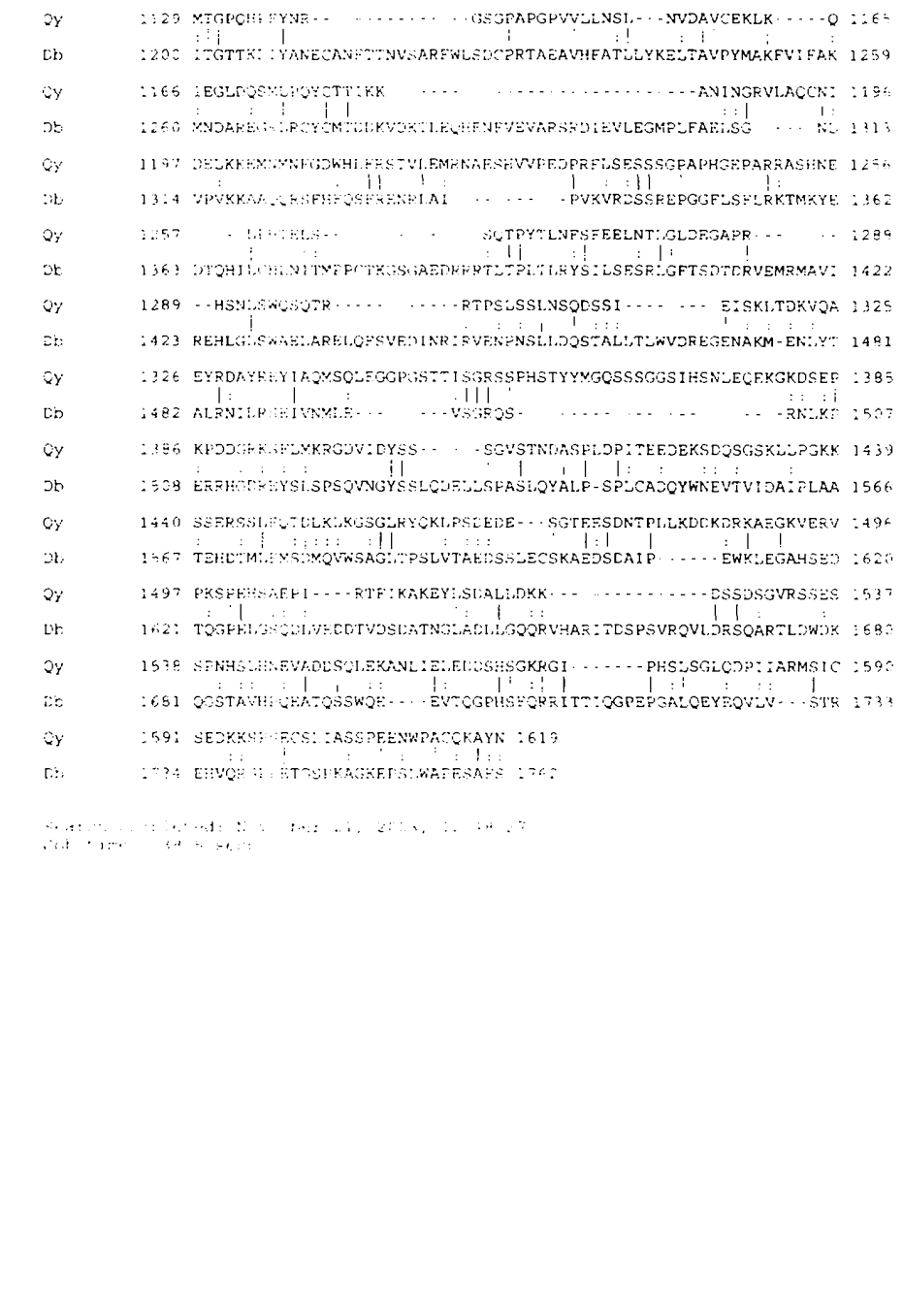




RESULT 1
S37771
ankyrin, erythrocyte - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: S37771
R;Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found
A;Reference number: S37771; MUID:93252825; PMID:8486643
A;Accession: S37771
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1848 <BIP>
A;Cross-references: EMBL:X69663; NID:g311816; PIDN:CAA48801.1; PID:g311817
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;48-80/Domain: ankyrin repeat homology <AN01>
F;81-113/Domain: ankyrin repeat homology <AN02>
F;114-146/Domain: ankyrin repeat homology <AN03>
F;147-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;275-307/Domain: ankyrin repeat homology <AN08>
F;308-340/Domain: ankyrin repeat homology <AN09>
F;341-373/Domain: ankyrin repeat homology <AN10>
F;374-406/Domain: ankyrin repeat homology <AN11>
F;407-439/Domain: ankyrin repeat homology <AN12>
F;440-472/Domain: ankyrin repeat homology <AN13>
F;473-505/Domain: ankyrin repeat homology <AN14>
F;506-536/Domain: ankyrin repeat homology <AN15>
F;[illegible]/Domain: ankyrin repeat homology <AN16>
F;[illegible]/Domain: ankyrin repeat homology <AN17>
F;[illegible]/Domain: ankyrin repeat homology <AN18>
F;[illegible]/Domain: ankyrin repeat homology <AN19>
F;[illegible]/Domain: ankyrin repeat homology <AN20>
F;[illegible]/Domain: ankyrin repeat homology <AN21>
F;737-769/Domain: ankyrin repeat homology <AN22>
F;770-802/Domain: ankyrin repeat homology <AN23>

Query Match 5.7%; Score 506; DB 2; Length 1848;
Best Local Similarity 19.4%; Pred. No. 5.2e-20;
Matches 378; Conservative 284; Mismatches 653; Indels 634; Gaps 93;





GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 2:14:4 ; Search time 15.5 Seconds
without alignments
527.377 Million cell updates/sec

Title: US-10-021-571-4
Perfect score: 4884
Sequence: 1 MSVLLSQSVIVYYFRNIA ... DPGAAeeLGLPeeeeYIL 1247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 121861 seqs, 47020775 residues

Total number of hits satisfying chosen parameters: 121861

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

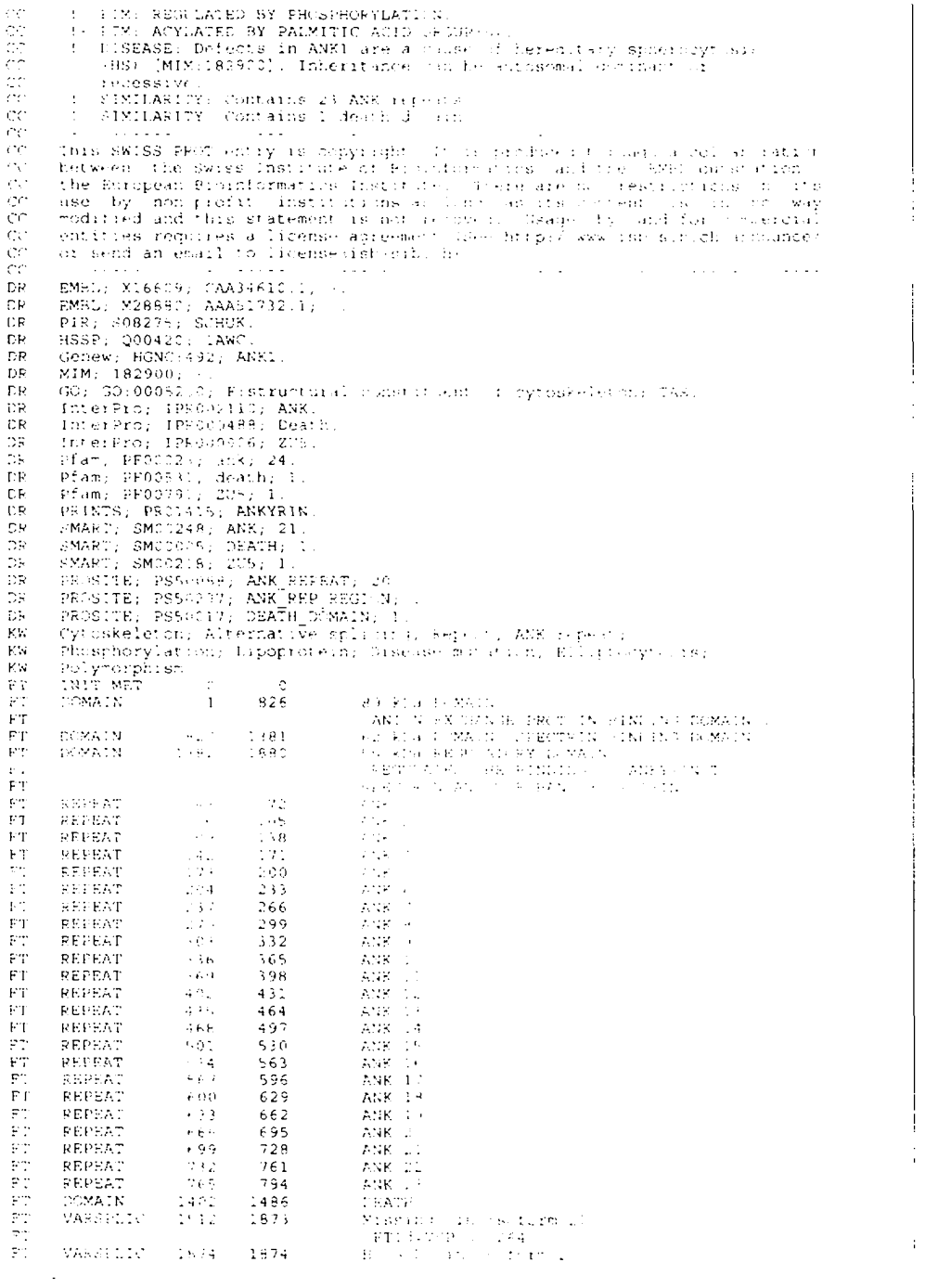
Database : SwissProt_41.*

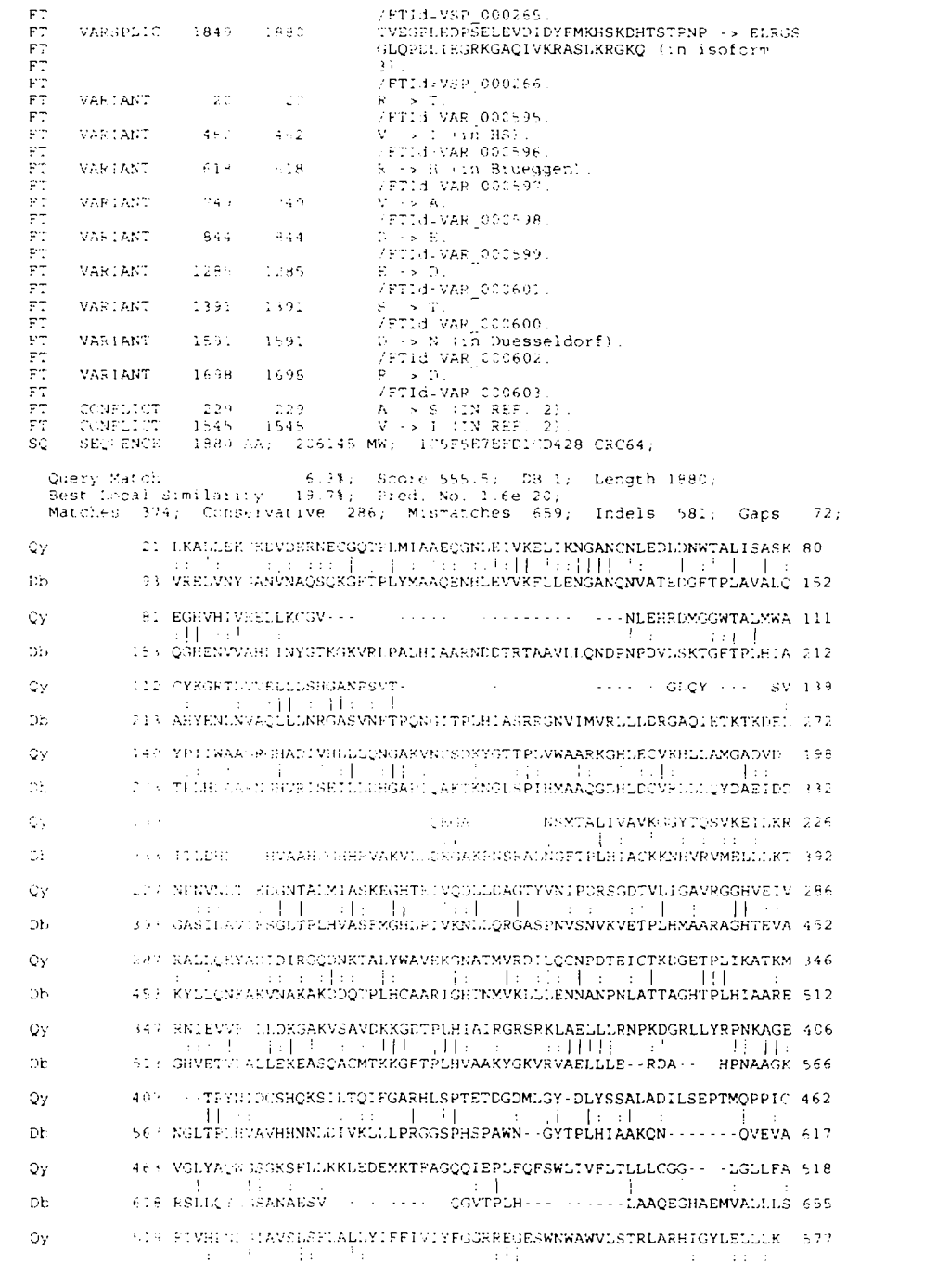
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distributions.

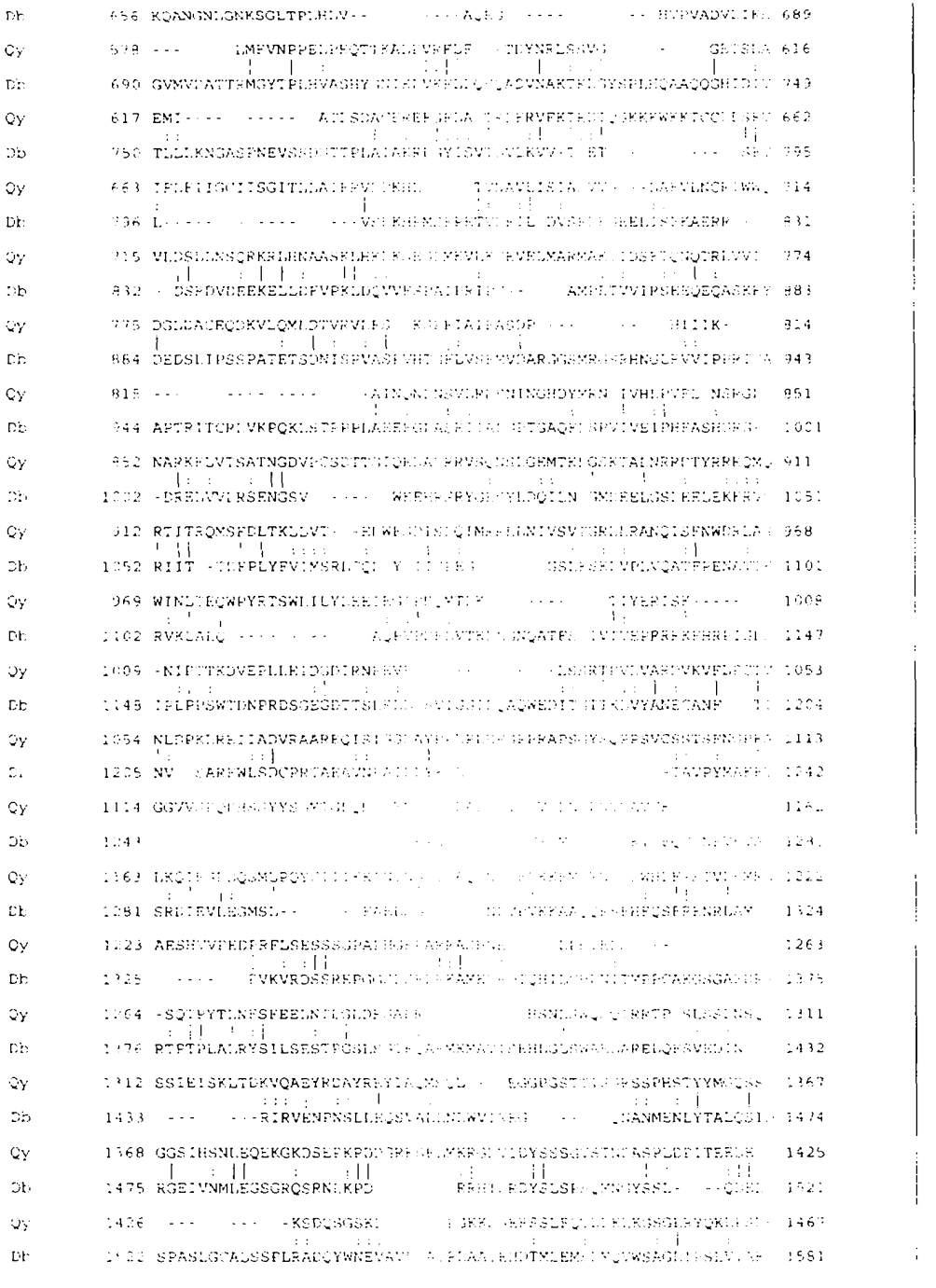
Result No.	Score	Query Match	Length	DB	ID	Description
1	555.5	11.4	1880	1	ANK1_HUMAN	P16157 homo sapien
2	539	11.1	3924	1	ANK2_HUMAN	Q12955 homo sapien
3	525	10.8	1862	1	ANK1_MOUSE	P54730 mus musculu
4	521	10.7	4377	1	ANK2_MOUSE	Q3UH54 homo sapien
5	458	9.4	768	1	ANK3_HUMAN	Q12955 homo sapien
6	448.5	9.2	1059	1	ANK3_MOUSE	Q8CHH9 homo sapien
7	429.5	8.8	1651	1	ANR1_HUMAN	Q15942 homo sapien
8	426.5	8.7	1427	1	ANR1_MOUSE	Q9Z2X1 homo sapien
9	376.5	7.7	423	1	ANR9_HUMAN	Q96BM9 homo sapien
10	373.5	7.6	987	1	ANR1_RAT	Q9JLM8 rattus norv
11	370	7.6	1166	1	ANR9_MOUSE	Q8K1M6 homo sapien
12	353	7.2	741	1	ANR49_HUMAN	Q8N8S7 homo sapien
13	351	7.2	692	1	ANR11_HUMAN	Q6UB99 homo sapien
14	346.5	7.1	832	1	ANR49_MOUSE	Q8BZ03 homo sapien
15	348	7.1	735	1	RN168_MOUSE	Q8BLL4 mus musculu
16	338	6.9	1267	1	HMIP_HUMAN	P49023 homo sapien
17	337.5	6.9	518	1	ASB4_HUMAN	Q9BYM8 homo sapien
18	328	6.7	1401	1	LATA_LATCA	Q27951 latimeria
19	324.5	6.7	525	1	ASB4_MOUSE	Q9JLM8 mus musculu
20	318.5	6.5	747	1	VD22_FOWPV	Q91951 fowlpox vir
21	318	6.5	583	1	AS11_MOUSE	Q8VBW4 mus musculu
22	309	6.3	596	1	VD26_FOWPV	Q91954 fowlpox vir
23	307	6.3	668	1	VD44_FOWPV	Q91A26 fowlpox vir
24	304.5	6.2	542	1	VD27_FOWPV	Q91954 fowlpox vir
25	302	6.2	436	1	VD40_FOWPV	Q91A40 fowlpox vir
26	284	5.8	1083	1	YIL2_YEAST	P41808 saccharomyc
27	280	5.7	1210	1	BAT3_HUMAN	P46379 homo sapien
28	279	5.7	1263	1	BAT3_MOUSE	Q91YA6 mus musculu
29	278	5.7	525	1	VD29_FOWPV	Q91954 fowlpox vir
30	275	5.6	592	1	VD34_FOWPV	Q91A34 fowlpox vir
31	267	5.5	603	1	VD41_FOWPV	Q91A41 fowlpox vir
32	264.5	5.4	656	1	PEX1_YEAST	P32793 saccharomyc
33	264.5	5.4	1150	1	ANK1_PIG	P29744 sus scrofa
34	259.5	5.3	1134	1	ANS1_HUMAN	Q92625 homo sapien
35	259	5.3	433	1	AS14_MOUSE	Q8VHS7 mus musculu
36	254.5	5.2	434	1	AS15_HUMAN	Q8WXK1 homo sapien
37	252	5.2	318	1	ASB7_MACFA	Q9BGT9 macaca fasc
38	252	5.2	642	1	YA2A_SCHPO	Q09701 schizosacch
39	247.5	5.1	318	1	ASB7_MOUSE	Q91ZU0 mus musculu
40	244.5	5.0	334	1	ASB7_HUMAN	Q9H672 homo sapien
41	242.5	5.0	429	1	AS10_HUMAN	Q8WXI3 homo sapien
42	240.5	4.9	278	1	AS13_HUMAN	Q8WXK3 homo sapien
43	240.5	4.9	776	1	ANR6_HUMAN	Q9NU02 homo sapien
44	239	4.9	434	1	V023_FOWPV	Q9J5B8 fowlpox vir
45	239	4.9	333	1	ANR2_HUMAN	Q9GZV1 homo sapien

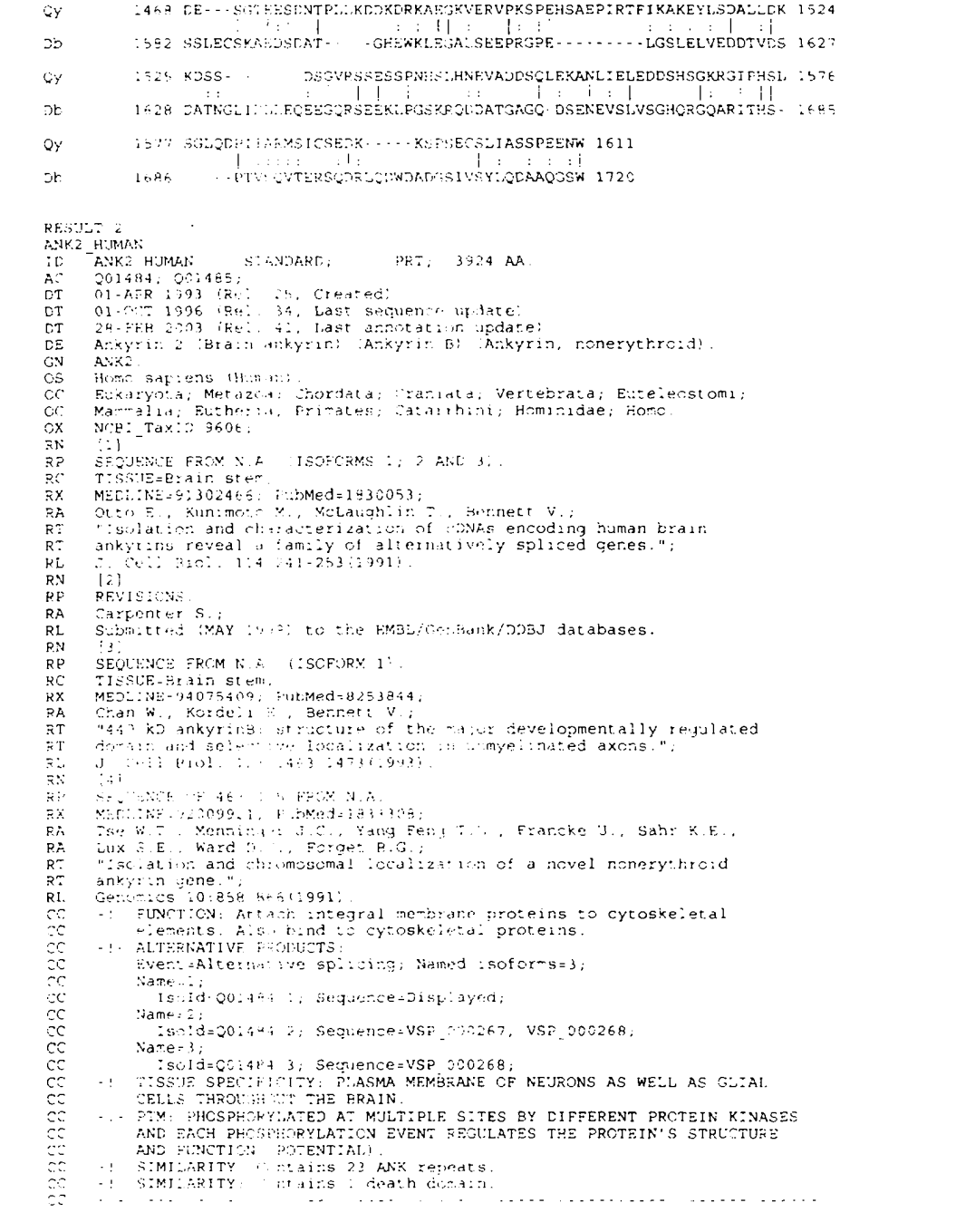
ALIGNMENTS

RESULT 1
ANK1_HUMAN
ID ANK1_HUMAN STANDARD; PRT; 1880 AA.
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP 2003 (Rel. 42, Last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE=Hematopoietic;
RX MEDLINE=90158811; PubMed=2110557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell cycle
RT control proteins.";
RL Nature 344:36-42 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175870; PubMed=1689849;
RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
RA Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN [3]
RP VARIANT HS ILE 9.
RX MEDLINE=94062594; PubMed=7504251;
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA Dornwell M., Herbers J., Kugler W., Ozcan R., Pekrun A.,
RA Gallagher P.G., Schroter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive
RT hereditary spherocytosis.";
RL Nat. Genet. 8:211-218(1994).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; PINS TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN SOME OR ALL OF THESE BINDING FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA
CC MEMBRANE.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=2.1;
CC Isoid=P16157-1; Sequence Displayed;
CC Name=2; Synonyms=2.2;
CC Isoid=P16157-2; Sequence=VSP_000264, VSP_000265;
CC Name=3;
CC Isoid=P16157-3; Sequence=VSP_000266;









```
RESULT 2
ANK2_HUMAN
ID   ANK2_HUMAN      STANDARD;      PRT;  3924 AA.
AC   Q01484; Q01485;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN   ANK2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).
RC   TISSUE=Brain stem.
RX   MEDLINE=91302466; PubMed=1930053;
RA   Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT   "Isolation and characterization of cDNAs encoding human brain
RT   ankyrins reveal a family of alternatively spliced genes.";
RL   J. Cell Biol. 114:241-253(1991).
RN   [2]
RP   REVISIONS.
RA   Carpenter S.;
RL   Submitted (MAY 1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain stem.
RX   MEDLINE=94075409; PubMed=8253844;
RA   Chan W., Kordeli E., Bennett V.;
RT   "440 kD ankyrinB: structure of the major developmentally regulated
RT   domain and selective localization in unmyelinated axons.";
RL   J. Cell Biol. 123:1463-1473(1993).
RN   [4]
RP   SEQUENCE OF 46-[illegible] FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Tse W.T., Menninger J.C., Yang Feng T.L., Francke U., Sahr K.E.,
RA   Lux S.E., Ward D.C., Forget B.G.;
RT   "Isolation and chromosomal localization of a novel nonerythroid
RT   ankyrin gene.";
RL   Genomics 10:858-866(1991).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements. Also bind to cytoskeletal proteins.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=Q01484-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q01484-2; Sequence=VSP_000267, VSP_000268;
CC       Name=3;
CC         IsoId=Q01484-3; Sequence=VSP_000268;
CC   -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC       CELLS THROUGHOUT THE BRAIN.
CC   -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC       AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC       AND FUNCTION (POTENTIAL).
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   -!- ----------- ---- - - -- ---- - - - ----- ---------- ------ ------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -!- ---------- ---
DR   EMBL; X56657; CAA40278.1.
DR   EMBL; X56658; CAA40279.1.
DR   EMBL; Z26634; CAB46644.1.
DR   EMBL; M77123; AAA50828.1.
DR   PIR; S37431; S37431.
DR   HSSP; P42771; 1IOC.
DR   Genew; HGNC:493; ANK2.
DR   MIM; 106410;
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF02791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 24.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 24.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Phosphorylation.
FT   REPEAT      63     92       ANK
FT   REPEAT      96    125       ANK
FT   REPEAT     129    158       ANK 3
FT   REPEAT     162    191       ANK 4
FT   REPEAT     195    220       ANK
FT   REPEAT     231    261       ANK 6
FT   REPEAT     265    294       ANK
FT   REPEAT     298    327       ANK 8
FT   REPEAT     331    360       ANK 9
FT   REPEAT     364    393       ANK
FT   REPEAT     397    426       ANK 11
FT   REPEAT     430    459       ANK 12
FT   REPEAT     464    492       ANK 13
FT   REPEAT     496    525       ANK 14
FT   REPEAT     529    558       ANK 15
FT   REPEAT     562    591       ANK 16
FT   REPEAT     595    624       ANK 17
FT   REPEAT     628    657       ANK 18
FT   REPEAT     661    690       ANK 19
FT   REPEAT     694    723       ANK 20
FT   REPEAT     727    756       ANK 21
FT   REPEAT     760    789       ANK 22
FT   REPEAT     793    822       ANK 23
FT   DOMAIN    1778   1960       [illegible]
FT   REPEAT    1778   1784       REPEAT A
FT   REPEAT    1796   1796       REPEAT A
FT   REPEAT    1797   1808       REPEAT A
FT   REPEAT    1809   1820       REPEAT A
FT   REPEAT    1821   1832       REPEAT A
FT   REPEAT    1833   1844       REPEAT A
FT   REPEAT    1845   1856       REPEAT A
FT   REPEAT    1857   1867       REPEAT A  A.T.R KIMATH
FT   REPEAT    1868   1879       REPEAT A
FT   REPEAT    1880   1891       REPEAT A
FT   REPEAT    1892   1902       REPEAT A  ATTR KIMATH
FT   REPEAT    1903   1914       REPEAT A
FT   REPEAT    1915   1926       REPEAT A
FT   REPEAT    1927   1938       REPEAT A
FT   REPEAT    1939   1950       REPEAT A
FT   DOMAIN    3536   3620       DEATH
FT   VARSPLIC  1039   1039       Q -> [illegible]
FT                               /FTId=VSP_000267.
FT   VARSPLIC  1444   1528       Missing (in isoform 2 and isoform 3).
FT                               /FTId=VSP_000268.
FT   CONFLICT   471    476       GQ -> PE (IN REF. 4).
FT   CONFLICT   971    971       I -> S (IN REF. 1).
FT   CONFLICT  3581   3582       QY -> HA (IN REF. 1).
FT   CONFLICT  3586   3586       I -> Y (IN REF. 1).
SQ   SEQUENCE  3924 AA;  440[illegible] MW;  5AC496C428E29D2 CRC64;

  Query Match              6.1%;  Score 543;  DB 1;  Length 3924;
  Best Local Similarity    20.1%;  Pred. No. 3.5e-19;
  Matches 421;  Conservative  294;  Mismatches  717;  Indels  662;  Gaps  41;

Qy          14 RRENIPA-AILERGFSVIERNPTSQTPLMIAAEQGNLEIVKELIKNGANCNLRDLFNWT 73
               :|  :    :|:  :  |   :  :  |  :  .|  :|  :| ||  :  : :|
Db          73 KEGHVSI-LLLGRGSHVLSATKRGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFT 132

Qy          74 ALISANKP-EGDHIVEELLKCGVN-  --  ---        ---------- LEHRDMG--- 113
               |  :::    :  :|  ||: | |                          :  :
Db         133 PLYMAAQE-NHIDVVKYLLENGANQNTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVR 192

Qy         104 ----       - ----------    -    - ------GWTALMWACYKGRTDVVELL 153
                                                        |:  .  |  :|  :|  ||
Db         193 LPALHIAARDDTKSAALLLQNDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLL 252

Qy         154 LSHGANINVTGDQYSVYPIIWAAGRGHAPIVHLLLQNGAKVNCSDKYGTTPLVWAARKG 213
               |: :|    |  :  :  :  |: :: |||  |  :::   :  |   |  |:. |
Db         253 LNRGAAVDIFA RNGITPLHVASKEGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 312

Qy         214 -          - ---  - - -    -- --HLECVKHLLAMGADVDQEGANSMTALIV 270
                                                |:: |||:   | ||:    : :::  |
Db         313 DQVVELLIKRGAPLLARTKNGLSPLHMAAQGDHVECVKHLLQHKAPVDDVTLDYLTALHV 372

Qy         271 AVKGGYTDVVKEILKRNPNVNLTDKDGNTALMIASKE------  ------------ -- 312
               |   |:       :| :  | |    :|   | ||  :
Db         373 AAHCGHYEVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGASIQAITESG 432

Qy         313 ----       GHTEIVQDLLDAGTYVNIRDRSGDTVLIGAVRGGHVEIVRALLQKYADID 362
                      |  ||  :|  :  :: :  |::   |    :||: :|:  :  :|
Db         433 LTPIHVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCLLRNGALVD 492

Qy         363 IRGQDNKTALYWAVEKGNATMVRDILQQNPDTEICTKDGETPLIKATKMRNIEVVELLLD 422
               :  :|:  :  :  :  :|  :  :|  :  :| |:|  :  :  :|  :  :
Db         493 ARARECTPLHIASRLGKTEIVQLLLQHKAHPDAATTNGYTPLHISAREGQVDVASVLLE 552

Qy         423 KGAKVNAK-MRDETPLHIAIRGRSEPLAELLLR--      -    --- - ----NPKD 474
                  :  :|  :  :  :  :|  :                            |
Db         553 A-AAA-    -RETPLHVAAKYGH  DAELLLQRHAAADSAGKNGLTPLHVAAHYDNQKV 611

Qy         475 --  RGLL       YRFNKASETPYNIA  AKLKSILTQI  EGAP       -  --HLSPT 522
                     |         :  |  |:  :          :  :|              :  :
Db         612 ALLLLDR-  L PHATARNGYTPLHIAARHNQMQIASTLLNYGAETNIVTKQGVTPLHLASQ 671

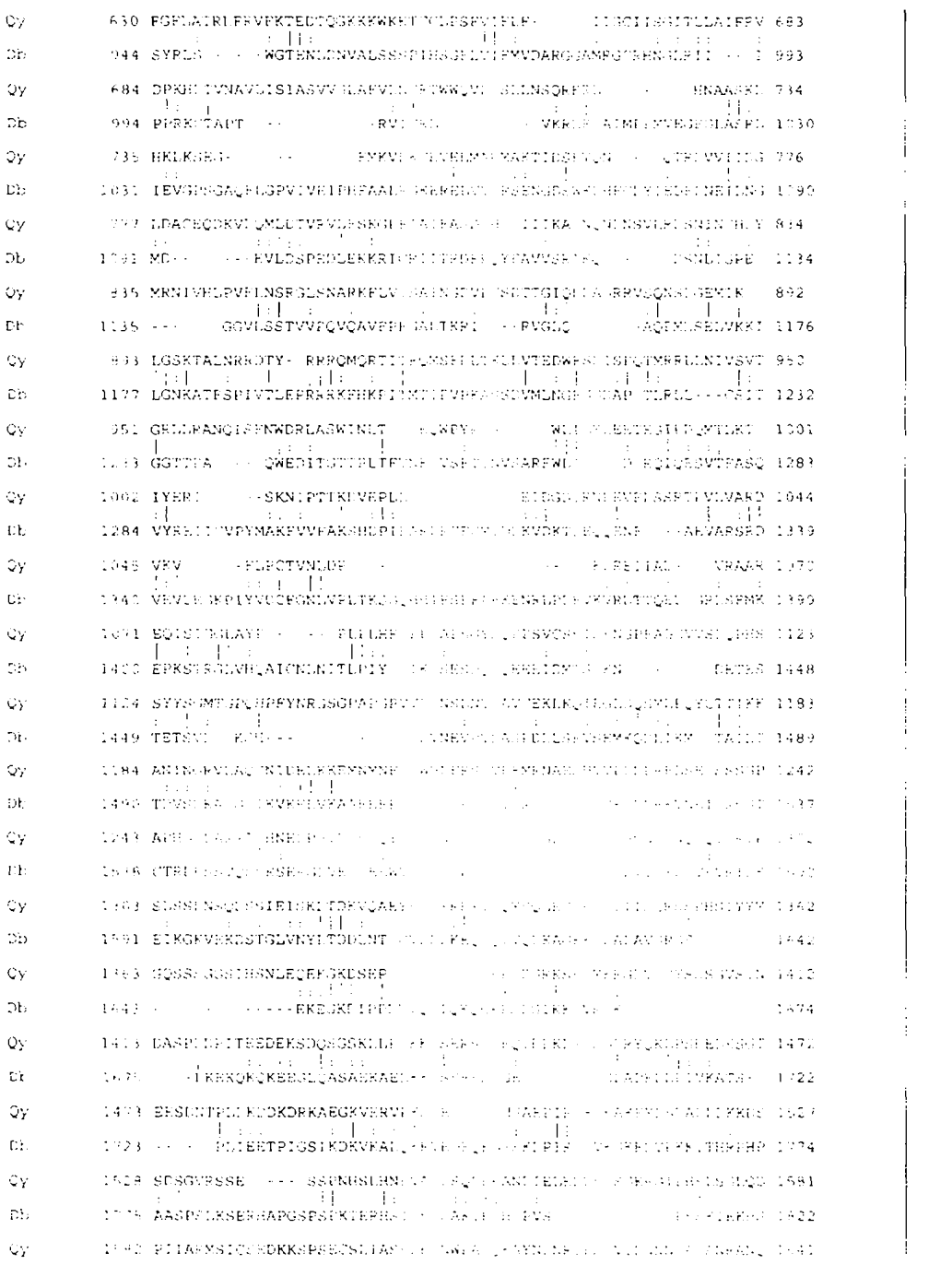
Qy         523 RTDGDYI     --GYDLYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKLEDEM - 484
                  |         |  :  |  :  :          :  :|              ||::
Db         672 EGHTDMVSLLLQKGANIHMSTKSGLTS  ------LHLAAQ-----  -    --EDKVNV 719

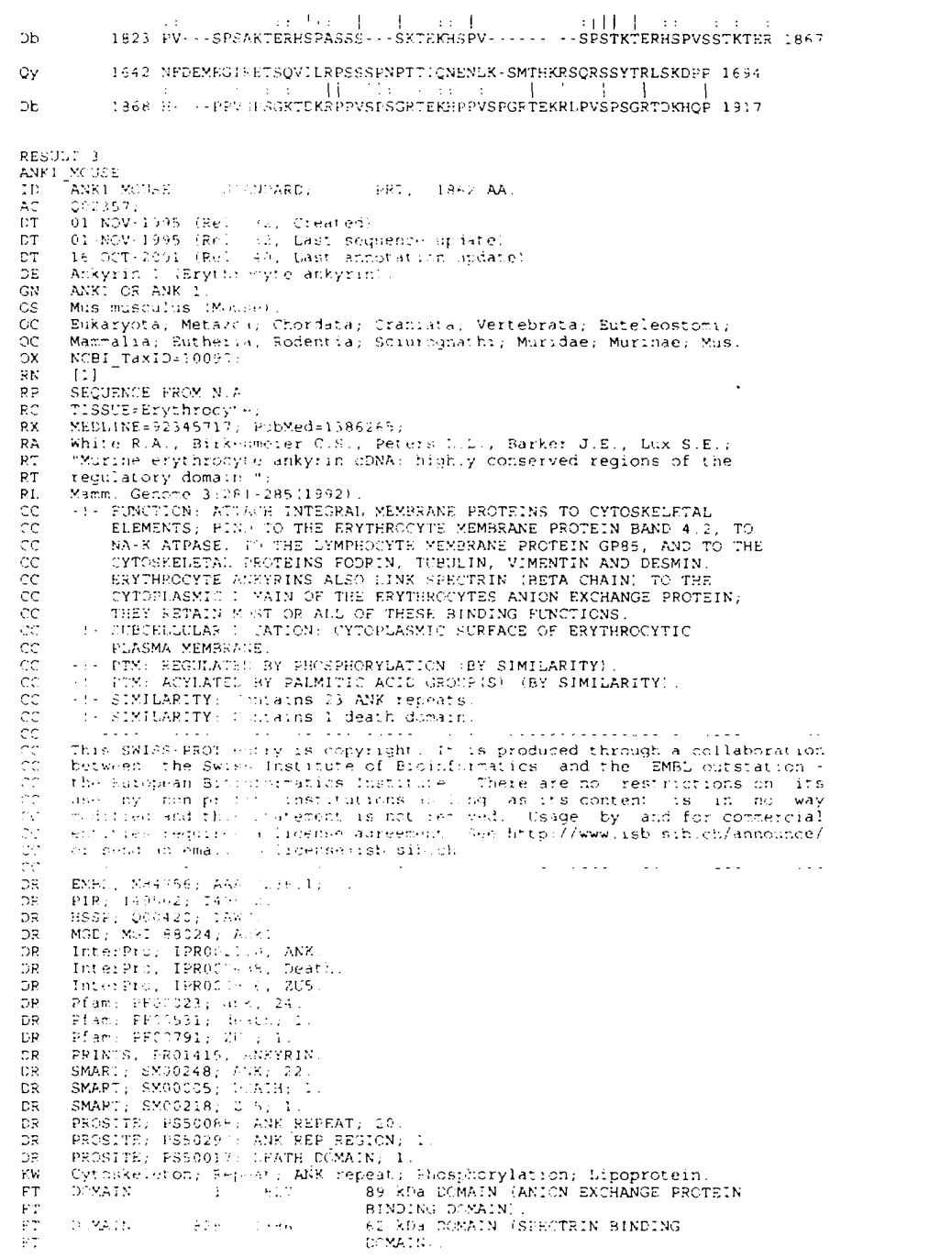
Qy         485     --RTLAQQIEPLFQFSWLIVFLTLLLC--GGLGLD-FAFTVHPNLGIAVSLSFLAL 539
                       :  :||      |  :  :  |       |:          :  :
Db         720 ADILTHGAIQDAHTKLGYTPLIV ----ACHYGNVKMVNFLLKQGANVNAKTKNGYTPL 775

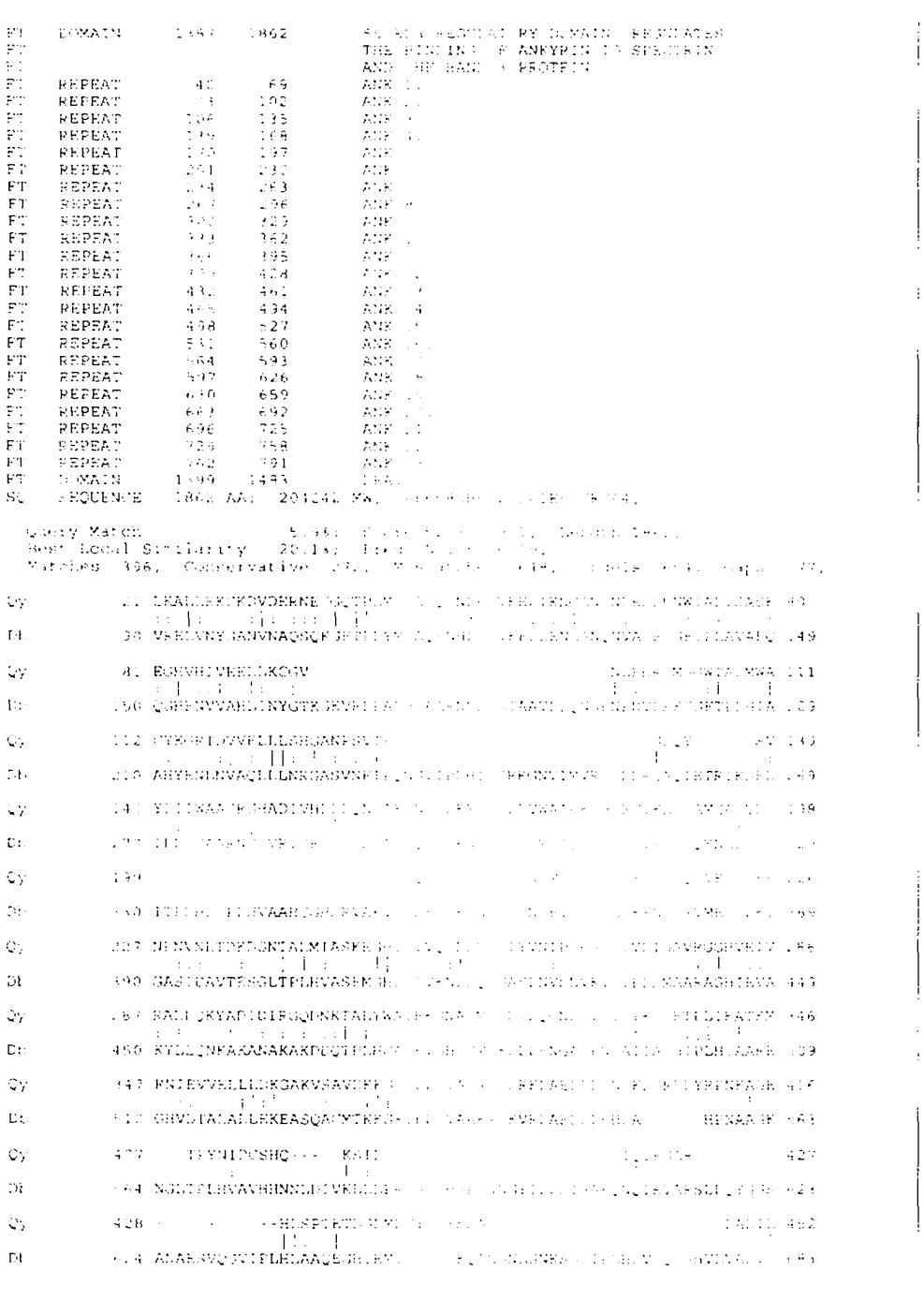
Qy         540 --       DYIPPIVIYFGGPREGESWNWAWVLSTRLARHIGYLELL--LKLM------- 579
                        :|  :::  :  :  |    |:  :|: :| : ::  ||:
Db         776 HQAAQQGHTHIINVLLQHGAKPNATTANGNTALA--IAKRLGYISVVDTLKVVTEEVTTT 833

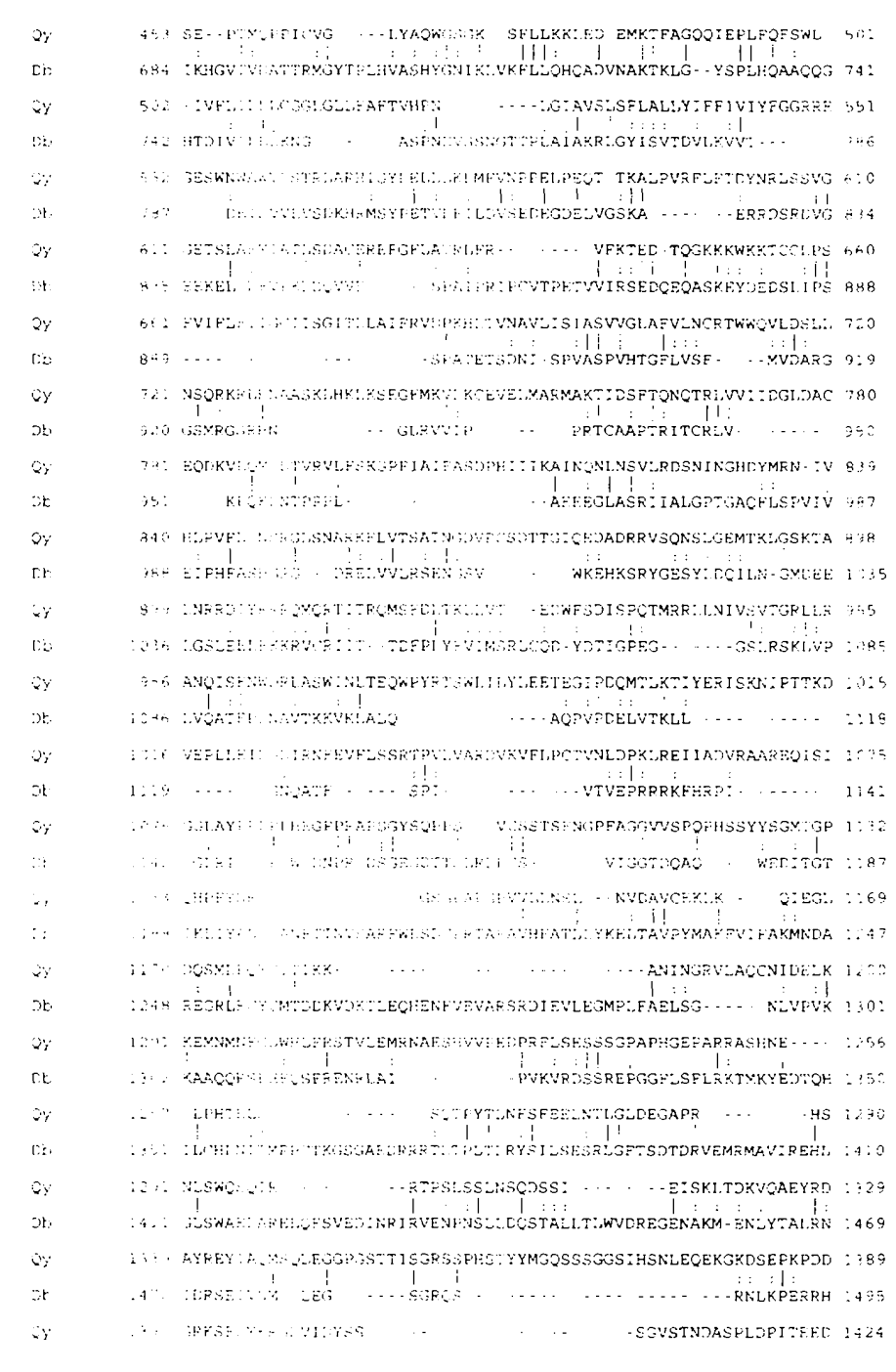
Qy         580 -----  -  ---------------- --  --------FVNPPELPEQTTKALP-VRFL-F 620
                                                    :: | :| |    :||  : !
Db         834 TTTITEKN-KLNVPETMTEVLEVSDERGDDTMTGDGGEYLRPEDLKELGDDSLPSSQFLDG 893

Qy         621 TDYNPLS-VGGET-SL--------       ----------------AEMIATLSDACERE 669
                  :|  :  :  :  :::                            :  :|||:  :
Db         894 MNYLRYPL-GGRSDSLRSFSSDRSHILSHASYLNDSAVMDDSVVIPSHQVSTLAKEAERN 953
```









```
Db       1496 GDREYSLSPSQVNGYSSLQDELDSTASILQVAIDVSRLCADQYRNEVTVIDAIFL AATERI 1554

Qy       1425  --EKSDQS -GSKLLPGKKSSERGSLPQT      - - DLPLKSSGIPYQKIF  - SP 1466

Db       1555 TMLPYSDMQVWSAGLTPSLVTAEDASILSKAIHDAIPEWKLETAHSEITDIFELSQI 1614

Qy       1467 SDESITEESDNT  --PLDRDLKDSFAFIKVEEVPKSPEHLAPIIATPIPAPGYLLDAI 1522

Db       1615 LVEDMCVDSDATNGLASLLSQRRRLYNPPPKLIV SSTEQLSPVSIVSLLEVHAPII 1673

Qy       1523 DKK---   DSSDSGVRNSESSPNHCTPEVAT QIEKAN  PIPILAGDFPGI     1572

Db       1674 DSPSVFQVLDRSQAETLDWTKQDSIAVAIQEAILSAQE    KPIQSIPLPRIIPLV 1729

Qy       1573 ---PRSLSGEQDPIIARMSICSEDFEVSFEIDIIASSPEENMPAIQKAYN 1619

Db       1730 GFEDSAICEYEQVLV---STREHVQPDVETLIKAGKEPSLWAIESAFS 1776


RESULT 4
ANK3_HUMAN
ID   ANK3_HUMAN     STANDARD;      PRT:  4377 AA.
AC   Q12955;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin 3 (ANK 3) (Ankyrin G).
GN   ANK3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain stem;
RX   MEDLINE=95138209; PubMed=7836469;
RA   Kordeli E., Lambert S., Bennett V.
RT   "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RT   at the axonal initial segment and node of Ranvier.";
RL   J. Biol. Chem. 270:2352-2359(1995).
CC   -!- FUNCTION: Membrane-cytoskeleton linker.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=A number of isoforms are produced.
CC       Name=1;
CC         IsoId=Q12955-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q12955-2; Sequence=VSP_000744;
CC       Name=3;
CC         IsoId=Q12955-3; Sequence=VSP_000745;
CC     -!- TISSUE SPECIFICITY: Expressed in brain, heart, skeletal muscle,
CC     -!- SIMILARITY: Contains 24 ANK repeats.
CC     -!- SIMILARITY: Contains 1 death domain.
CC     -!- SIMILARITY: Contains 1 SPEC repeat.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U13616; AAA64834.1; -.
DR   PIR; A55575; A55575.
DR   HSSP; P55273; 1RI8.
DR   Genew; HGNC:494; ANK3.
DR   MIM; 600465; -.
DR   GO; GO:0006605; P:protein targeting; NAS
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT   REPEAT       73    102       ANK 1.
FT   REPEAT      106    135       ANK 2.
FT   REPEAT      139    168       ANK 3.
FT   REPEAT      172    201       ANK 4.
FT   REPEAT      203    230       ANK 5.
FT   REPEAT      234    263       ANK 6.
FT   REPEAT      267    296       ANK 7.
FT   REPEAT      301    329       ANK 8.
FT   REPEAT      333    362       ANK 9.
FT   REPEAT      366    395       ANK 10.
FT   REPEAT      399    428       ANK 11.
FT   REPEAT      432    461       ANK 12.
FT   REPEAT      465    494       ANK 13.
FT   REPEAT      498    527       ANK 14.
FT   REPEAT      531    560       ANK 15.
FT   REPEAT      564    593       ANK 16.
FT   REPEAT      597    626       ANK 17.
FT   REPEAT      630    659       ANK 18.
FT   REPEAT      663    692       ANK 19.
FT   REPEAT      696    725       ANK 20.
FT   REPEAT      729    758       ANK 21.
FT   REPEAT      762    791       ANK 22.
FT   REPEAT      795    825       ANK 23.
FT   DOMAIN     1519   1898       SER-RICH.
FT   DOMAIN     4090   4174       DEATH.
SQ   SEQUENCE   4377 AA;  480399 MW;  F4237985576883684 CRC64;

  Query Match             6.9%;  Score 921;  DB 1;  Length 4377;
  Best Local Similarity   20.6%;  Pred. No. 4.5e-18;
  Matches  427;  Conservative  284;  Mismatches  695;  Indels  668;  Gaps   87;

Qy         14 EEENIPALKALLEKGKDVDERNECGQTPLMIAAEQGNLEIVKELIKNGANCNLEDLDNWT 73

Db         84 XEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFT 142

Qy         74 ALISASKKEGVVHIVEELLKCGVNLEHRDMGGWTALMWACYKGRTDVVELLL------ 124

Db        143 PLYMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLQENDTKGKVR 202

Qy        125 ---   -     --      -          - ---  SHGANPSVTGL  ---  -- 135

Db        203 LPAGHIAV--FDSTKAAAILQNDNMADVESPSGFTPLHIAAHYGNINVATLLLNRAAAVD 262

Qy        136     QYDVSQTLWNAAGRGNADIVSHLLQNGAPVNCSDKYGTTPLVWAARKG-    --- -  182

Db        263 RTARNGYTPLHVASKEGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLD 322

Qy        183  -  --      ----     ----HLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQS 219

Db        323 RAAPILSKTKGGLSPLHMATQGDHLNCVQLLLQHNVPVDDVTNDYLTALHVAAHCGHYKV 382

Qy        220 VKEILKRNPFNVNLTDKDGNTALMIASKE---------------------- ----- - 247

Db        383 AKVLLTKFANPNAKALNGFTPLHIACKKNRIKVMELLLKHGASIQAVTESGLTPIHVAAF 442

Qy        248 -GHTEIVQLLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRGQDNKTA 306

Db        443 MGHVNIVSQLMHHGASPNTCNVRGETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTP 502

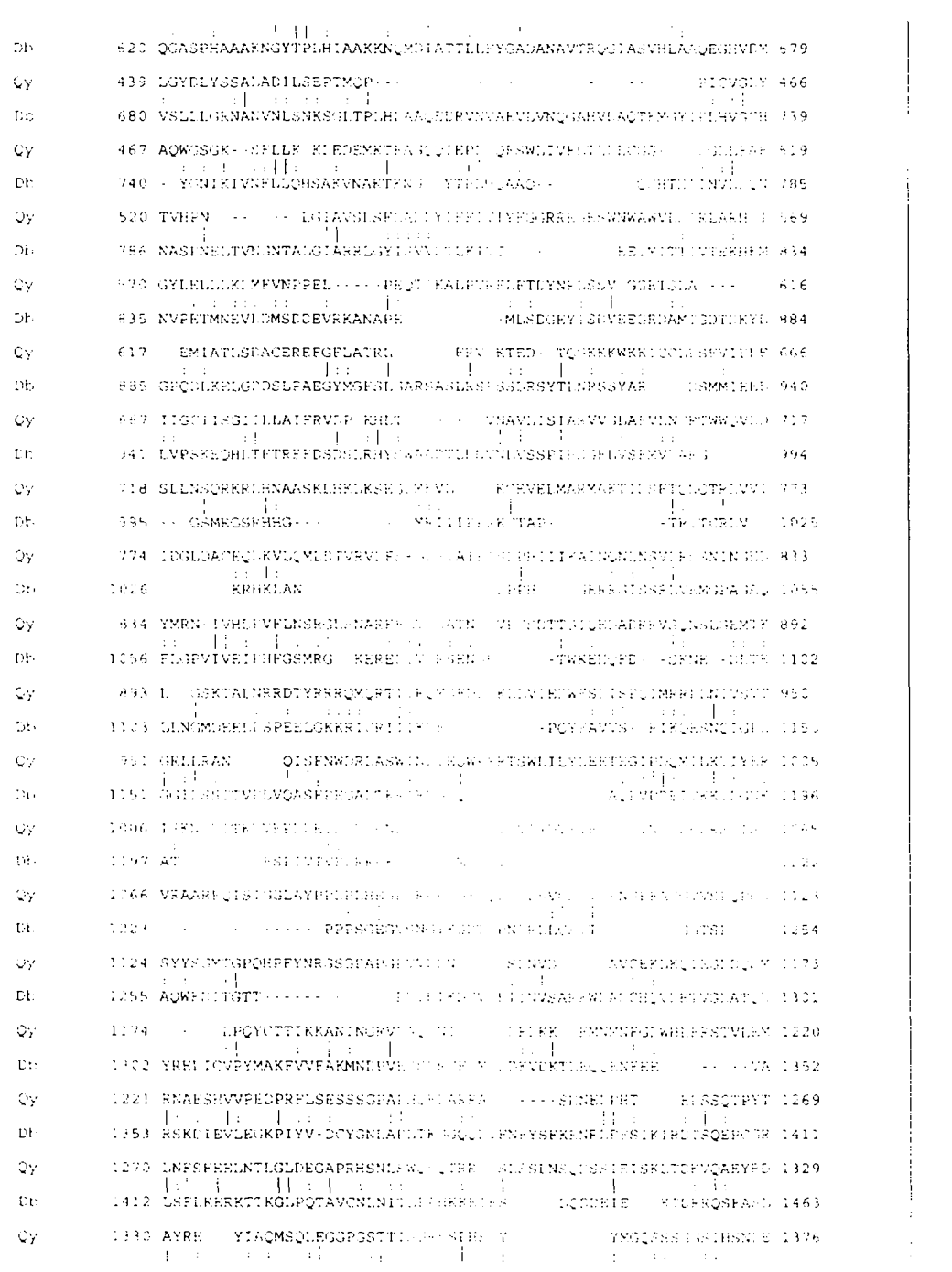
Qy        307 LYWAVEKQNATMVRDILQ--CNPDT-- ---------------------------EICT 333

Db        503 LHISARLGKADIVQQLLQQGASPNAATTSGYTPLHLSAREGHEDVAAFLLDHGASLSITT 562

Qy        334 KDGETPLHKATKYRNIEVVELLLDKGAKVSAVDKKGDTPLHIAIRGRSRKLAELLLRNPK 393

Db        563 KKGFTPLQCAAKYGKLEVANLLLQKGASPDAAGKSGLTPLHVAAHYDNQKVALLLL- -D 619

Qy        394 DGRLDVPLKKAGRTPYDLDD-SHQKSILTQI--PGAP----------HLSHTETDGDM 438
```



```
Db        620 QGASPHAAAARNGYTPLHIAAKKNQMDIATTLLEYGADANAVTRQGIASVHLAAQEGHVDM 679

Qy        439 LGYDLYSSALADILSEPTMQP--                                PICVSLY 466

Db        680 VSLILGRNADVNLSNKSGLTPLHLAAQEGRVDVAEVLVNQGAHVEAQTRMGYTPLHVATH 739

Qy        467 AQWGSGK--GFLLP KIEDEMKTRAFQGIRPI QRSWLIVELIIGTGD--  GDLEAP 519

Db        740 - YGNIKIVNFLLQHSAPVNAKTPKG  YTPLQLAAQ--          QIETTIINVLLQY 795

Qy        520 TVHFN  --   -- LGIAVGLSPLAIIYIPFLIYFLIYPGRRA-RNWWAAVIL PLAAH I 569

Db        796 NASINELTVEGNTALGIARRLGYISVVLVLPTLI  -          EELYITPIIYEKHFM 854

Qy        570 GYLELLLKLMFVNPPEL-----PEQT KALPTAFLETLYNFLSDV GGELGLA ---    616

Db        855 NVPETMNEVIGMSDDEVRKANAPE          -MLSDGRYIGDVEESDDAMIIDDICKYL 904

Qy        617   EMIATLSPACEREFGFLATRL      PPS KTED- TQIKKKWKKIGNLGMVIEIF 666

Db        905 GPQDLKRLGNDSLFAEGYMGESLIARNASLANISSCRSYTIRPSGVAP    DSMMIEEE 960

Qy        667 IIGTIIGGIILLAIPRVDP KRIT    - -  VMAVLISIASVV NAFVLN PTWWLVLI 717

Db        961 LVPSKRQHLTPTREEDSDDLRHVGAATTLIMCNLVSSPIIYGLVSERVV A  S       994

Qy        718 SLLNSQRKRLHNAASKLHFQKSEIYRVL    RTEVELMARVAFTIIPFQLQTPLVVI 773

Db        995 -- GSMRQSWHHG---         - YSIIIIIKTTAP-          -TPLTIRIV   1026

Qy        774 IDGLDAFEQLKVSLMLDTVRVLFA- - -AE- -QIPIII-PAIQNDNLNSVIE SNIN EG 833

Db       1026         KRHKLAN                  IPEP    IAAAIIDSFIDSMFFAMLL 1080

Qy        834 YMRN-IVHLIVFLNSMGLLNAFF- - -ADN  VE YGTTIIQRGAEFFVSLGSLEMTP 892

Db       1081 FLSPVIVEIIHFGSMRG  KERELV SGEN -      -TWKEDQFD  -GFNR -GITR 1132

Qy        893 L  GSKIALNRRDTYRRRQMVRTIFAMVFVG KLNYIEFWFSIISFLIMARNLINIVSGT 952

Db       1133 LLNGMIRELISPEELGKKRIGWIIFVT S          -PQTVAVDS- PIPQAPNQIGLL 1192

Qy        953 GRLLRAN     QIGFNWDRIASWIK-RGWK-PTGWLILPLLEEATGGIFQQVIDKIIYRF 1006

Db       1193 GQIISGITVELVQASPFEQAICE-PV- -  _              ALIVPTFIAKLIPF 1246

Qy       1007 LFD- -TFKYDPIFIL-  - -VL  - -  - -PVLVP-  -  -  -VL  - - -V- - 1055

Db       1247 AT        RSEIVVIL---  -  -  -  -  -  -  -  -  -  -  -  -  -  1282

Qy       1056 VRAARLVSIGSLAYPFCFIMGA- V-- -- -  - -VL  - -SVRPVVPGIVGLI- - 1114

Db       1283 -  -  -  -  ----- PPFSVSPVVPIFKGT PVPKIDG-I          IGSI     1334

Qy       1115 SVYSLYTGPQHFFYNRSGPAPGIAIGW  SIGVD  AVTGFLKQLLSIGLV 1173

Db       1335 AQWFIITGTT------ -  -  -  -  -  -  -  -  -  -  -  -  -  -  1360

Qy       1174 -  LPQYCTTIKKANIMGRVI-  -  -  -  -  -  -  -  -  -  -  -  -  1220

Db       1361 YRELICVEYMAKFVVFAKMNDPVL-  -  -  -  -  -  -  -  -  -  -  -  -  1402

Qy       1221 RNAESHVVPEDPRPLSESSSGPAILGMIAPA  ----SINEIHT  EIASQTPYT 1269

Db       1403 RSKDIEVLEGKPIYV-DCYGNLAFLTA-  -  -  -  -  -  -  -  -  -  -  -  1411

Qy       1270 LNFSFEELNTLGLPEGAPRHSNLSWL-  -  -  -  -  -  -  -  -  -  -  -  1329

Db       1412 DSFLKERKTTKGLPQTAVCNLNIL-  -  -  -  -  -  -  -  -  -  -  -  -  1463

Qy       1330 AYRE   YIAQMSQLEGGPDSTTI-  -  -  -  -  -  -  -  -  -  -  -  -  1376

Db       1464 ALRKRYSYLTEPGMIERSTGAT----RSLP TTYSYKPFFSTRPYQSWTTAPITV----- 1513

Qy       1377 QEKGKDSPPKPDDGRKSFLMKRGDVIDYSSSGVSTNDASPLDPI----TEEDEKSDQSGS 1432

Db       1514 -------  -PGPAKSGF-  -  -TSLSSSSSNTPSASPLKSIWSVSTPSPIKSTLGAS 1556

Qy       1433 KLLPGKKSERSSLFQIDLKLKGSGLR---YQKLPSDEDESGT-----EESDNTPL---- 1480

Db       1557 CTSSVFSIGDVASPIRS-LRTMSSPIKTVVSQSPYNIQVSSGTLARAPAVTPATPLKGLA 1615

Qy       1481  - -- LPIDKDRKAEGKVER      VPKSPE -----HSAEPIRTFI KAKEYLSDAL 1521

Db       1616 SNSTPSGISPVTTAGSLLERSSITYIPIASPKSNINMYSSSLPFKSIITSAAPLISSPL 1675

Qy       1522  ----     -IDKKDSSDSGVRSSESSPNHSL--HNEVA----DDSQLEKAN---LIE-- 1560

Db       1676 KSVVSPVFSPVDVISSAKITMASSLSSPVKQMPGHAEVALVNGSISPLKYASSSTLINGC 1735

Qy       1561      -LEIDSHSGKRGIPHSLSGIQLPTIARMSICSEDKKSPSECSLIASSPEENWPACQ 1615

Db       1736 KATATLQRKISSATNSVSSVVSAATDTVEKVFSTTTAMPFSPLR SYVSAAP  - - -  1786

Qy       1616 KAYNLNRIISTVTLNNNSAPANRANQNPDEMEGIRETSQVILRPSSS--   PNPTTIQN 1670

Db       1787 SAPQSLRTISASALYTSLGSSISATTS     ---SVTSSIITVPVYSVVNVLIEPALKKL 1839

Qy       1671 ENLKSMTHRESQRSS- YTRLSKDPPELHAAASS 1702

Db       1840 PDSNSFTGAAALLSPIKTLTTETHPQPEFSRTS 1873


RESULT 5
YB23_HUMAN
ID   YB23_HUMAN     STANDARD;      PRT;   768 AA.
AC   Q9ULJ7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein KIAA1223 (Fragment).
GN   KIAA1223.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20239461; PubMed=10774462;
RA   Nagase T., Ishikawa K., Kikuno R., Hirosawa M., Nomura N.;
RA   Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
CC   --- SIMILARITY: Contains at least 14 ANK repeats.
CC   -!- ---------- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.com).
CC   ---- ---- ----- ---------- ---------- ---------- ---------- ---- -------
DR   EMBL; AB033049; BAA86537.1; -.
DR   HSSP; P42771; 1DCQ.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 14.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
```

```
FT   NON_TER       1      1
FT   REPEAT        2     31       ANK 1.
FT   REPEAT       16     44       ANK 2.
FT   REPEAT       48     82       ANK 3.
FT   REPEAT       86    115       ANK 4.
FT   REPEAT      119    148       ANK
FT   REPEAT      152    181       ANK
FT   REPEAT      185    214       ANK
FT   REPEAT      218    247       ANK
FT   REPEAT      251    280       ANK
FT   REPEAT      284    313       ANK
FT   REPEAT      317    346       ANK
FT   REPEAT      350    379       ANK
FT   REPEAT      383    412       ANK
FT   REPEAT      416    446       ANK 28
SQ   SEQUENCE   768 AA;  82819 MW;  [illegible]

  Query Match             5.2%;  Score 41.5;  DB 1;  Length 768;
  Best Local Similarity   27.9%;  Pred. No. 4.1e-06;
  Matches  122;  Conservative   78;  Mismatches  163;  Indels   74,  Gaps    4;

Qy         24 LLEKCKDVDERNECGQTPLMIAAE[illegible]IVKELIKNGANCNL 72
              |:  :|:: :  |:  : |:  :|  :  :  :|| |: | :|    |
Db          2 LLQHGARVNKADNEGRTALIAAAAYMGHIEIVEVLLKHGASVNIRETYGNTALMEAALTGH 61

Qy         73 ASKEGHVHIVEELLKQGVNLEHRDMFFWTALMWAVYKGRTDVVELLLRHGANPSVTNQLY 132
              ||'  |  :   |:   ::   :  '  |:  ::|  ,'|: :'  :::  |
Db         62 ASK--GHASVVSLLLQRGARVHHTTMYGITPLDLAAYFGHVEVVKFLLQAGAAVHHTD-EN 118

Qy        133 SVYPIIWAAQRGHADIVHLLI[illegible]                    150
             |::  |  '  :|:
Db        119 GRTPLLAAASMGHASVVNTLIFWGAAVNAVDSSGRTPVLSIAAAQGNVEVVRTLLDRGIDV 178

Qy        151 --   ----    --      -  [illegible]AKGRTDVYEKYGYTPLYWAAAFKNIEICEVLIL 191
                                      :  :  |  :  :  |  :  :  |  :  |
Db        179 NHRDDAGWTPIHMAAAFGHRDIVSYLIQQGADVNAQDNDGRIPIDLARQCGHDDIVSLLEE 238

Qy        192 AMGADVDQEGANSMTALIVAVF[illegible]KEIIKRGPNVNLIDKLGNTALMIASKFGHI 251
              :::  | :      ||  :  :  :  :  |  :  | :  :
Db        239 ENKSNIDQRGYDGRNALRVAAL[illegible]VELLIRHGADVNIKDMMGRIPLDLAIENGH 298

Qy        252 IVQELLDAGTYVNIRDRGGDTVLI[illegible]RRNPYVRALLQRYADIDIHQQDKRGALYWAI 311
              :  :  :| :  |  :  |  :  |  :  :  :  :  :  :  :  :  :|:|
Db        299 MAEYFLENGANVEASDAEGRTALH[illegible]MVQVLIAYHADVNAADNEKRSALHWA 358

Qy        312 EKGNATMVRFLQKNMDLTDYFF[illegible]KMDITEVDLLLKEKGAEVSAVDKFGKTPFDL 371
              :  :  :|:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db        359 WQGHEEVVQLLIEEGASTDAQN[illegible]NSALHLA 417

Qy        372 PDPEIKAILKDYKLKKLAQLAA 393
              :  :  :  :  :|:|
Db        418 AMRVAARNDHQGIIRVLQE 436


RESULT 6
Y379_HUMAN
ID   Y379_HUMAN     STANDARD;      PRT;  [illegible] AA.
AC   Q15084;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein KIAA0379 (Fragment).
GN   KIAA0379.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97349984; PubMed=9205841;
RA   Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.

RT   "Prediction of the coding sequences of unidentified human genes. VII.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 4:141-150(1997).
CC   -!- SIMILARITY: Contains at least 27 ANK repeats.
CC   -!- [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   [illegible]
DR   EMBL; AB002377; BAA20833.2; -.
DR   HSSF; P80144; DM00.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 28.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 28.
DR   PROSITE; PS50088; ANK_REPEAT; 24.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT       46     75       ANK 1.
FT   REPEAT       79    108       ANK 2.
FT   REPEAT      112    141       ANK 3.
FT   REPEAT      145    174       ANK 4.
FT   REPEAT      178    207       ANK 5.
FT   REPEAT      211    240       ANK 6.
FT   REPEAT      244    273       ANK 7.
FT   REPEAT      277    307       ANK 8.
FT   REPEAT      311    340       ANK 9.
FT   REPEAT      344    373       ANK 10.
FT   REPEAT      377    406       ANK 11.
FT   REPEAT      410    439       ANK 12.
FT   REPEAT      443    472       ANK 13.
FT   REPEAT      476    506       ANK 14.
FT   REPEAT      510    540       ANK 15.
FT   REPEAT      551    584       ANK 16.
FT   REPEAT      588    617       ANK 17.
FT   REPEAT      622    651       ANK 18.
FT   REPEAT      658    687       ANK 19.
FT   REPEAT      691    720       ANK 20.
FT   REPEAT      724    753       ANK 21.
FT   REPEAT      761    790       ANK 22.
FT   REPEAT      794    823       ANK 23.
FT   REPEAT      828    857       ANK 24.
FT   REPEAT      861    891       ANK 25.
FT   REPEAT      895    924       ANK 26.
FT   REPEAT      931    960       ANK 27.
SQ   SEQUENCE  1059 AA;  113465 MW;  C1F55E6CFE494770 CRC64;

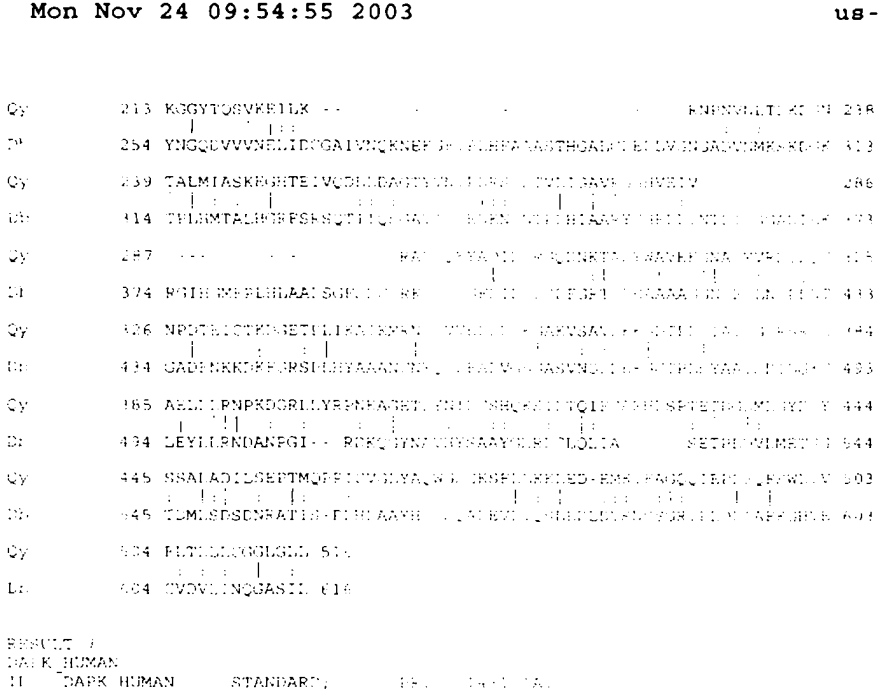
  Query Match             4.9%;  Score 438.5;  DB 1;  Length 1059;
  Best Local Similarity   25.4%;  Pred. No. 7e-15;
  Matches  156;  Conservative   95;  Mismatches  243;  Indels  119;  Gaps   13;

Qy          7 QSVINYVEEENIPALKALLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELIKNGANCNL 66
              ::: |  :|    :: : | : |: ::   :::  ||  |: |:: || :|:  |
Db         20 QAIFNGIISG--  VRALIFKKEDVNFQDNEKRTPLHAAAYLGDAEIIELLILSGARVNA 75

Qy         67 EDLDNW    -TALISASKEG-------  -------------HVHIVEELLKCG----  94
              :|  .      |: : |::                        |:    : :||
Db         76 KD-SKWLTPLHRAVASCSEEAVQVLLKHSADVNARDKNWQTPLHIAAANKAVKCAEALVP 134

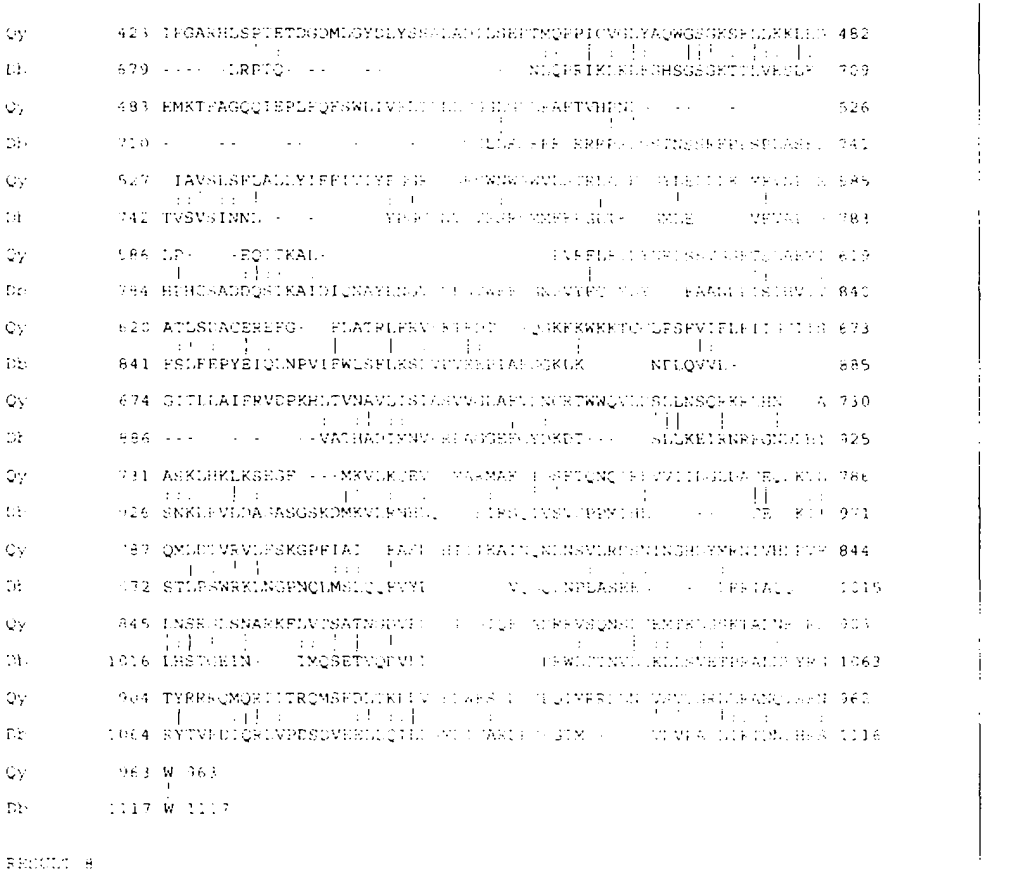
Qy         95 --VNLEHRDMGGWTALMWACYKGRTDVVELLLSHGANPSVTGLQYSVYPIIWAAGRGHAD 152
                :    |  :  |  :  |  :|:|:|| ||| :   :      '|| :  :
Db        135 LLSNVNVLDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFD-KKDRRAIHWAAYMGHIE 193

Qy        153 IVHLLIQKAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAV 212
              :  :||:   |:.  ||  |||  ||   :  ||: | :  |:::       | :  :
Db        194 VVKLLASHGASVTCKDKKSYTPLHAAAASGMISVVKYLLDLGVDMNEPNAYGNTPLHYAC 253
```

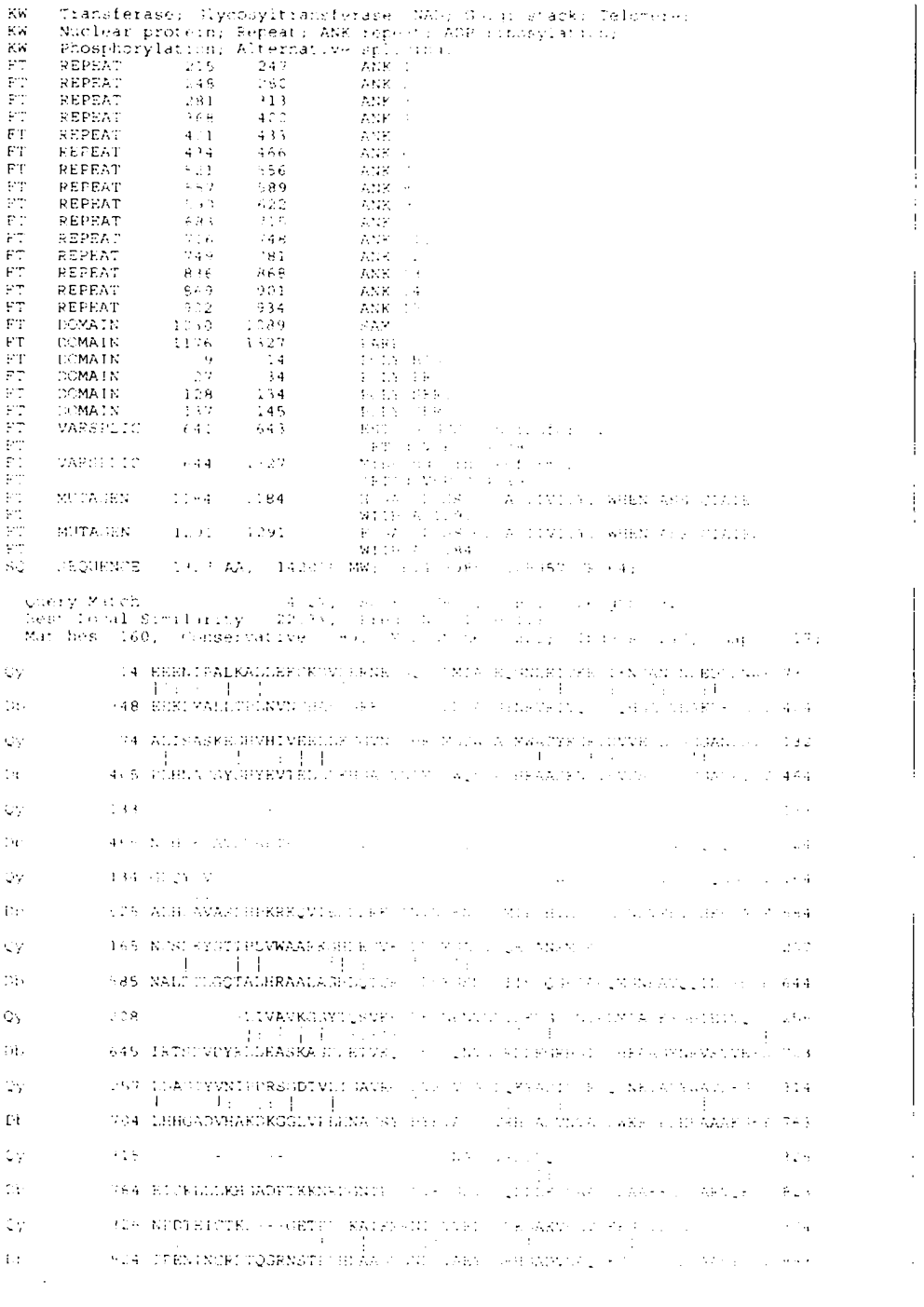


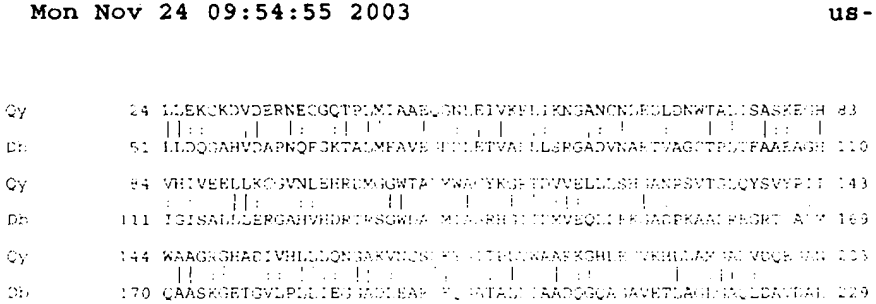
```
RESULT 1
DAPK_HUMAN
ID   DAPK_HUMAN     STANDARD;      [illegible] AA.
AC   P53355;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Death-associated protein kinase [illegible] (EC [illegible]).
GN   DAPK1 OR DAPK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; [illegible];
OC   Mammalia; Eutheria; Primates; [illegible]; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95[illegible];
RA   Deiss L.P., [illegible];
RC   "Identification [illegible]
RT   protein as [illegible]
RT   death.";
RL   Genes Dev. [illegible] (1995).
RN   [2]
RP   REVISIONS TO 164-171.
RA   Feinstein E.;
RL   Submitted (APR-1997) to the EMBL [illegible].
CC   -!- FUNCTION: INVOLVED IN MEDIATING [illegible]
CC       DEATH.
CC   -!- PTM: AUTOPHOSPHORYLATED.
CC   -!- SIMILARITY: BELONGS TO THE [illegible]
CC   -!- SIMILARITY: Contains 10 ANK [illegible]
CC   -!- SIMILARITY: Contains 1 death [illegible]
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are restrictions on its
CC   use by non-profit institutions as well as [illegible] way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X76104; CAA53712.1; -.
DR   PDB; 1JKT, 01-APR-02.
DR   GeneW; HGNC:2674, DAPK1.
DR   MIM; 600891; -.
DR   GO; GO:0004684, [illegible]calmodulin-dependent protein kinase I activity; TAS.
DR   GO; GO:0008624, [illegible]induction of apoptosis by extracellular sig. . .; TAS.
DR   GO; GO:0006468, [illegible]protein amino acid phosphorylation; TAS.
DR   InterPro; IPR0[illegible]; ANK.
DR   InterPro; IPR0[illegible]; Death.
DR   InterPro; IPR0[illegible]; Prot kinase.
DR   InterPro; IPR0[illegible]; Ser_Thr_pkinase.
DR   Pfam; PF00023, [illegible]; 4.
DR   Pfam; PF00531, [illegible]; 1.
DR   Pfam; PF00069, [illegible]; 1.
DR   ProDom; PD000[illegible]; Prot_kinase, 1.
DR   SMART; SM00248, [illegible]; 7.
DR   SMART; SM00005, [illegible]; 1.
DR   SMART; SM00220, [illegible]; 1.
DR   PROSITE; PS50[illegible]; ANK_REPEAT; 6.
DR   PROSITE; PS502[illegible]; ANK_REP_REGION; 1.
DR   PROSITE; PS500[illegible]; DEATH_DOMAIN; 1.
DR   PROSITE; PS00107, PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011, PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108, PROTEIN_KINASE_ST; 1.
KW   Transferase; Ser/Thr protein kinase; Calmodulin-binding;
KW   Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW   3D-structure.
FT   DOMAIN       [illegible]    266       PROTEIN KINASE.
FT   DOMAIN       267    334       CALMODULIN-BINDING.
FT   REPEAT       378    407       ANK 1.
FT   REPEAT       411    440       ANK 2.
FT   REPEAT       444    473       ANK 3.
FT   REPEAT       478    507       ANK 4.
FT   REPEAT       511    540       ANK 5.
FT   REPEAT       544    573       ANK 6.
FT   REPEAT       577    606       ANK 7.
FT   REPEAT       610    639       ANK 8.
FT   REPEAT       87[illegible]    705       ANK 9.
FT   REPEAT       116[illegible]   1197      ANK 10.
FT   DOMAIN       131[illegible]   1397      DEATH.
FT   NP_BIND       [illegible]     29        ATP (BY SIMILARITY).
FT   BINDING       42      42        ATP (BY SIMILARITY).
FT   ACT_SITE      139     139       BY SIMILARITY
FT   MUTAGEN       42      42        K >A; LOSS OF ACTIVITY.
SQ   SEQUENCE   1431 AA;  160017 MW;  9EF84811004A155B CRC64;

  Query Match             4.6%;  Score 469.5;  DB 1;  Length 1431;
  Best Local Similarity   [illegible]%;  Pred. No. 3.4e-33;
  Matches [illegible]; Conservative 149;  Mismatches 138;  Indels  [illegible];  Gaps  36;
```



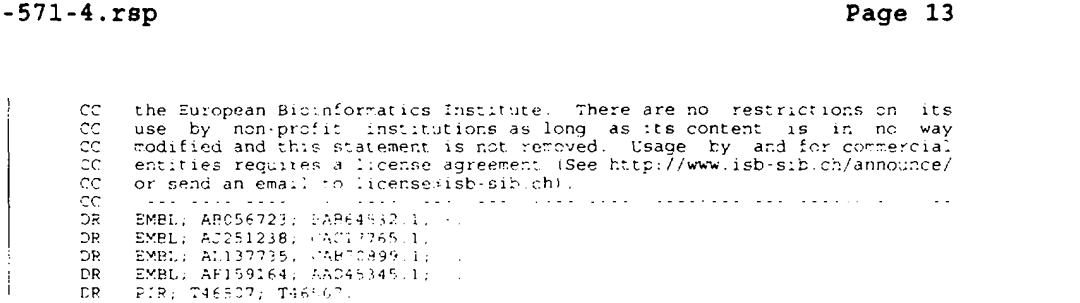
```
RESULT 8
TNK1_HUMAN
ID   TNK1_HUMAN              STANDARD;      PRT;   1327 AA.
AC   O95271; Q96J23;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase-1) (TNKL) (TNK) (Ser
DE   interacting ankyrin-related ADP ribose polymerase).
GN   TNKS OR TNKS1 OR TIN1 OR TINF1 OR PARP5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Testis;
RX   MEDLINE=99040105; PubMed=9822378;
RA   Smith S., Giriat I., Schmitt A., de Lange T.;
RT   "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
RL   Science 282:1484-1487(1998).
RN   [2]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=99454782; PubMed=10523313;
RA   Smith S., de Lange T.;
RT   "Cell cycle dependent localization of a telomeric PARP.";
RT   to nuclear pore complexes and
RL   J. Cell Sci. 112 (Pt 21):3649-3656(1999).
RN   [3]
RP   FUNCTION, AND PHOSPHORYLATION.
RX   MEDLINE=20556262; PubMed=10988299;
RA   Chi N.W., Lodish H.F.;
RT   "Tankyrase is a golgi associated mitogen-activated protein kinase
RT   substrate that interacts with IRAP in GLUT4 vesicles.";
RL   J. Biol. Chem. 275:38437-38444(2000).
RN   [4]
RP   FUNCTION, AND MUTAGENESIS OF HIS 1194 AND GLU 1291.
RX   MEDLINE=21602814; PubMed=11719545;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC       subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
CC       the cytoplasmic domain of LNPEP/Otase in SLC2A4/GLUT4-vesicles.
CC       Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4 vesicles. A minor proportion is
CC       also found at nuclear pore complexes and around the pericentriolar
CC       matrix of mitotic centromeres. During interphase, a small fraction
CC       of TNKS is found in the nucleus, associated with TRF1.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         Isoid=O95271-1; Sequence=Displayed;
CC       Name=2;
CC         Isoid=O95271-2; Sequence=VSP_004538, VSP_004539;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC   -!- PTM: Upon insulin stimulation, phosphorylated on serine residues
CC       by MAPK kinases.
CC   -!- PTM: ADP-ribosylated (auto).
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: Contains 25 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---------- ---- ---- ---- -------- ---- ---- ---- ---- ---- ----- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---- ---- -------- ---- ---- ---- ---- ---- ----- --
DR   EMBL; AF082556; AAC79841.1; -.
DR   EMBL; AF082557; AAC79842.1; -.
DR   EMBL; AF082558; AAC79843.1; -.
DR   EMBL; AF082559; AAC79844.1; -.
DR   HSSP; Q00420; 1AXB.
DR   Genew; HGNC:11941; TNKS.
DR   MIM; 603303; -.
DR   GO; GO:0005696; C:telomere; IDA.
DR   GO; GO:0003950; F:NAD+ ADP-ribosyltransferase activity; IDA.
DR   GO; GO:0005515; F:protein binding activity; IPI.
DR   GO; GO:0007004; P:telomerase-dependent telomere maintenance; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 19.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 17.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
```


KW Transferase; Glycosyltransferase; DNA; G-rich; Track; Telomere;
KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
KW Phosphorylation; Alternative splicing.
FT REPEAT 215 247 ANK 1
FT REPEAT 248 280 ANK 2
FT REPEAT 281 313 ANK 3
FT REPEAT 368 402 ANK 4
FT REPEAT 411 435 ANK
FT REPEAT 474 466 ANK 6
FT REPEAT 521 556 ANK 7
FT REPEAT 557 589 ANK 8
FT REPEAT 593 622 ANK 9
FT REPEAT 686 715 ANK
FT REPEAT 716 748 ANK 11
FT REPEAT 749 781 ANK 12
FT REPEAT 816 868 ANK 13
FT REPEAT 869 901 ANK 14
FT REPEAT 902 934 ANK 15
FT DOMAIN 1060 1089 ADP
FT DOMAIN 1106 1327 ADP
FT DOMAIN 9 14 [10]
FT DOMAIN 27 34 [10]
FT DOMAIN 128 134 [10]
FT DOMAIN 137 145 [10]
FT VARSPLIC 641 643 MDC
FT PT
FT VARSPLIC 644 1327 MDC
FT PT
FT MUTAGEN 1184 1184 A
FT WT
FT MUTAGEN 1291 1291 A
FT WT
SQ SEQUENCE 1447 AA; 1620
Query Match
Best Local Similarity
Matches 160, Conservative
Qy 14 EEEN
Db 448 EGEG
Qy 74 ALI
Db 485 NALC
Qy 138
Db 495
Qy 186
Db 525 AGE
Qy 169 N
Db 585 NALC
Qy 208 -
Db 645 IATC
Qy 259 I
Db 704 LHHG
Qy 315
Db 764
Qy 408
Db 824
Qy 375 -AIRGRSRKLAELLLRNPKDGRLLYRPNKAGETP 408
Db 884 DIAALTPYNLVNAIDKWAFTPLHFAAQKGRT QICALLLAHGADPTM --KNQEGQTP 939
Qy 409 YN-- --IDSSHQKSILT --QIFGARHLSPTETDGDMLGYDLYSSALADI 451
Db 940 LDLATA--IPALLIDAMPFEALPTCFFPQATVVSASLISPASTPS ---CLSAASSIDN 994
Qy 452 LSEPTM-- PTYGIYAQWGGKKSFLKK -----LEDEMKTFAGQQIEPLFQ 493
Db 995 LTGPLAP LAVDGASNAGNAAASTERKEIEVAGLDMNISQFLKSLGLRHLRDIFE 1048
RESULT 9
ANKH_CHRVI
ID ANKH_CHRVI STANDARD; PRT; 323 AA.
AC Q06527;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ankyrin homolog precursor.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Chromatiaceae; Allochromatium.
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 9899 / DSM 180;
RX MEDLINE=93300844; PubMed=8390599;
RA Dolata M.M., van Beeumen J.J., Ambler R.P., Meyer T.E.,
RA Cusanovich M.A.;
RT "Nucleotide sequence of the heme subunit of flavocytochrome c from
RT the purple photosynthetic bacterium, Chromatium vinosum. A 2.6-kilobase
RT pair DNA fragment contains two multiheme cytochromes, a flavoprotein,
RT and a homolog of human ankyrin.";
RL J. Biol. Chem. 268:14426-14435(1993).
CC -!- SIMILARITY: Contains 9 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; L11544; AAA21997.1;
DR PIR; A41477; A41477.
DR HSSP; P02247; 1MBA.
DR InterPro; IPR002110; ANK
DR Pfam; PF00023; ank; 8.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 8.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Signal; ANK repeat; Repeat.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 323 ANKYRIN HOMOLOG.
FT REPEAT 31 60 ANK 1
FT REPEAT 64 93 ANK 2
FT REPEAT 97 126 ANK 3
FT REPEAT 130 159 ANK 4
FT REPEAT 163 192 ANK 5
FT REPEAT 196 225 ANK 6
FT REPEAT 229 258 ANK 7
FT REPEAT 262 291 ANK 8
SQ SEQUENCE 323 AA; 34494 MW; 3F7B1B4A9BD607DF CRC64;
Query Match 4.1%; Score 475.5; DB 1; Length 323;
Best Local Similarity 45.3%; Pred. No. 2e-12;
Matches 128; Conservative 49; Mismatches 128; Indels 1; Gaps 1;



RESULT 10
ASB2_HUMAN
ID ASB2_HUMAN STANDARD; PRT; 587 AA.
AC Q96Q27; Q9NSU5; Q9Y567;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin repeat and SOCS box containing protein 2 Asb-2.
GN ASB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21450456; PubMed=11566181;
RA Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yamori N.,
RA Nakanishi T., Tanaka K.;
RT "ATRA regulated Asb-2 gene induces differentiation of HL-60
RT leukemia cells.";
RL FEBS Lett. 508:223-228(2001).
RN [2]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21634875; PubMed=11690464;
RA Guibal F.C., Moog-Lutz C., Smolewski P., Di Cristofano A.,
RA Darzynkiewicz Z., Lutz P.G., Cayre Y.E.;
RT "ASB-2 inhibits growth and promotes differentiation of myeloid
RT cells.";
RL J. Biol. Chem. 277:218-224(2002).
RN [3]
RP SEQUENCE OF 31-587 FROM N.A.
RC TISSUE=Uterus;
RA Gottenwaelder B., Obermaier B., Mewes H.-W., Hansen A.,
RA Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE OF 42-587 FROM N.A.
RX MEDLINE=20564172; PubMed=11111046;
RA Kile B.T., Viney E.M., Willson T.A., Prodromidi E.I.,
RA Berlinguer A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA Alexander W.S.;
RT "Cloning and characterization of two mouse genes encoding new ankyrin repeat
RT and SOCS box-containing proteins Ask 1, Ask 2, Ask 3 and Ask 4.";
RL Gene 258:31-41(2000).
CC -!- FUNCTION: Not known. Its expression induces growth inhibition and
CC chromatin condensation.
CC -!- INDUCTION: By all-trans retinoic acid (ATRA).
CC -!- SIMILARITY: Contains 12 ANK repeats.
CC -!- SIMILARITY: Contains 1 SOCS box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB056723; BAB64852.1.
DR EMBL; AJ251238; CAC17765.1.
DR EMBL; AL137735; CAB70499.1.
DR EMBL; AF159164; AAD46345.1.
DR PIR; T46507; T46507.
DR Genew; HGNC:16021; ASB2.
DR MIM; 605759;
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001496; SOCS.
DR Pfam; PF00023; ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50225; SOCS; 1.
KW ANK repeat; Repeat.
FT REPEAT 56 85 ANK 1.
FT REPEAT 89 119 ANK 2.
FT REPEAT 123 152 ANK 3.
FT REPEAT 156 185 ANK 4.
FT REPEAT 189 218 ANK 5.
FT REPEAT 222 251 ANK 6.
FT REPEAT 255 284 ANK 7.
FT REPEAT 288 317 ANK 8.
FT REPEAT 320 349 ANK 9.
FT REPEAT 362 391 ANK 10.
FT REPEAT 392 421 ANK 11.
FT REPEAT 428 456 ANK 12.
FT DOMAIN 533 587 SOCS BOX.
FT CONFLICT 224 224 MISSING (IN REF. 4).
FT CONFLICT 460 460 MISSING (IN REF. 4).
SQ SEQUENCE 587 AA; 66083 MW; 666A2B6010C4C1F9 CRC64;

Query Match 4.2%; Score 373.5; DB 1; Length 587;
Best Local Similarity 29.8%; Pred. No. 6.2e-12;
Matches 111; Conservative 71; Mismatches 163; Indels 27; Gaps 6;



RESULT 11

```
TNK2_HUMAN
ID   TNK2_HUMAN      STANDARD;      PRT;  1166 AA.
AC   Q9H2K2; Q9H8F2; Q9HAS4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tankyrase 2 (EC 2.4.2.30) (TANK2 [tankyrase II] (TNKS 2) (PPP1
DE   interacting ankyrin-related ADP-ribosylase) (Tankyrase-like
DE   protein) (Tankyrase related protein).
GN   TNKS2 OR TNK1 OR TANK2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=21072550; PubMed=11909949;
RA   Monz D., Munnia A., Comtesse N., Kaufmann M., Stoeckel W., Feldhaus A.,
RA   Glass B., Meese E.U.;
RT   "Novel tankyrase-related gene detected with monoclonal antibodies specific
RT   sera.";
RL   Clin. Cancer Res. 7:1119-1126(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast carcinoma;
RX   MEDLINE=21196733; PubMed=11286487;
RA   Kuimov A.N., Kuprash D.V., Perlaki S.V., Chernichenko E.K.,
RA   Scanlan M.J., Jongeneel C.V., Lazar V.A., Nedospasov S.A.;
RT   "Cloning and characterization of human tankyrase genes
RT   family.";
RL   Genes Immun. 2:52-55(2001).
RN   [3]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC   TISSUE=Liver;
RX   MEDLINE=21264473; PubMed=11278494;
RA   Lyons R.J., Deane R., Lynch D.K., Ikin R.J., Sanders K.M.,
RA   Pyre R.C., Sutherland G.R., Daly R.J.;
RT   "Identification of a novel human tankyrase binding protein that interacts
RT   with the adaptor protein Grb14.";
RL   J. Biol. Chem. 276:17172-17180(2001).
RN   [4]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC   TISSUE=Placenta;
RX   MEDLINE=21443728; PubMed=11454888;
RA   Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D.,
RA   Morin G.B., Yaswen P., Campisi J.;
RT   "TANK2, a new TRF1 associated poly(ADP-ribose) polymerase with a unique
RT   rapid induction in senescent human cells.";
RL   J. Biol. Chem. 276:35951-35955(2001).
RN   [5]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE OF N-TERMINUS, AND SUBCELLULAR
RP   LOCATION.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=21661461; PubMed=11747774;
RA   Sbodio J.I., Lodish H.F., Chi N.-W.;
RT   "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1
RT   (telomere repeat-binding factor 1) and IRAP (insulin-responsive
RT   aminopeptidase).";
RL   Biochem. J. 361:451-459(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Yin Y., Gelmann E.P.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Tracey A.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 838-1161 FROM N.A.
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Isono K., Hosoiri T., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   FUNCTION, AND ADP-RIBOSYLATION.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC       subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS. Interacts with the
CC       cytoplasmic domain of INPPL1/SHIP2 in SLC2A4/GLUT4-vesicles. Binds
CC       to the N-terminus of Grb14 and TRF1 with its ankyrin repeat
CC       region.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4 vesicles. Also found around the
CC       pericentriolar matrix of mitotic centromeres. During interphase, a
CC       small fraction of TNKS2 is found in the nucleus, associated with
CC       TRF1.
CC   -!- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,
CC       liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral
CC       blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,
CC       ovary, small intestine, colon, mammary gland, breast and breast
CC       carcinoma, and in common-type meningioma. Highly expressed in
CC       fetal liver, heart and brain.
CC   -!- PTM: ADP-ribosylated (auto).
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: Contains 15 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ---------- ----------
DR   EMBL; AF305081; AAG25674.1; ALT_INIT.
DR   EMBL; AF449122; AAG44694.1; -.
DR   EMBL; AF329694; AAK11469.1; -.
DR   EMBL; AF442950; AAF29811.1; -.
DR   EMBL; AF409033; AAF42172.1; -.
DR   EMBL; AF438201; AAL40795.1; -.
DR   EMBL; AL359707; CAC07867.1; -.
DR   EMBL; AK023746; BAB14669.1; ALT_INIT.
DR   HSSP; Q02420; 1AWP.
DR   GeneW; HGNC:15677; TNKS2.
DR   MIM; 607128; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; Ank; 10.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.
FT   REPEAT       57     89       ANK 1.
FT   REPEAT       90    122       ANK 2.
FT   REPEAT      123    155       ANK 3.
FT   REPEAT      210    242       ANK 4.
```

```
FT   REPEAT      243    275       ANK [illegible]
FT   REPEAT      276    308       ANK [illegible]
FT   REPEAT      363    395       ANK [illegible]
FT   REPEAT      399    431       ANK [illegible]
FT   REPEAT      432    464       ANK [illegible]
FT   REPEAT      525    557       ANK
FT   REPEAT      558    590       ANK [illegible]
FT   REPEAT      591    62[illegible]  ANK [illegible]
FT   REPEAT      678    710       ANK [illegible]
FT   REPEAT      711    743       ANK [illegible]
FT   REPEAT      744    776       ANK [illegible]
FT   DOMAIN      8[illegible]   936    [illegible]
FT   DOMAIN      1[illegible]   1162   [illegible]
FT   CONFLICT    331    337       [illegible] VAA [illegible] IN REF [illegible]
FT   CONFLICT    357    361       NFK[illegible] IN REF [illegible]
FT   CONFLICT    966    966       Q [illegible] IN REF 8[illegible]
SQ   SEQUENCE    1166 AA;  1[illegible]6917 MW;  [illegible] [illegible]EF704 [illegible];

  Query Match             4.2%;  Score [illegible];  DB 1;  Length 1166;
  Best Local Similarity   25.6%;  Pred. No. [illegible];
  Matches  143;  Conservative  [illegible];  Mismatches  169;  Indels  [illegible];  Gaps  14;

Qy         19 PALKALLEKCKD[illegible]TPLMIAARQGN[illegible]VVELIKN[illegible]A 62
Db         23 PAARELFEACRN[illegible]VERVFSLVI[illegible]HPAAS[illegible]VAYLL[illegible]A 82

Qy         63 NCNLEDLDNWTALISASKEGHVH[illegible]WIALMWA[illegible]YK[illegible]RTDVV[illegible] 122
Db         83 NVCARDDGGLIPLHNACSFGHAEV[illegible]LEHR[illegible]NAFDNWNYT[illegible]HEAA[illegible]KIP[illegible] 142

Qy        123 LLSH[illegible]ANPSV[illegible] [illegible]QY 137
Db        143 LLQHGAEPTIRNTDGRTALDLA[illegible] [illegible]2

Qy        138 SVY[illegible] PI[illegible]WAAGF[illegible]HAD[illegible] [illegible]
Db        203 NCHA[illegible]DGRKSTPLHLAAGYNRVK[illegible]AFDE[illegible]HY[illegible] [illegible]

Qy        172 [illegible] TTPLVWAAR [illegible]AV[illegible]AD[illegible] [illegible]ALLVA 211
Db        263 LVKGHACVNAMDLWQFTPLHEAAS[illegible]SYGAD[illegible]HNF[illegible]AIDLA[illegible] 322

Qy        212 [illegible]VKG-GYTQSVKE [illegible] [illegible]
Db        323 KERL[illegible]YEFE[illegible]HSL[illegible]AAPF[illegible]VIE [illegible] 382

Qy        2[illegible] [illegible] [illegible]
Db        383 [illegible]ANINEK[illegible]KEF[illegible]HV [illegible] 442

Qy        280 G[illegible]HVE[illegible]IVRALLQKYADIDIRGQDN[illegible] [illegible] 327
Db        443 CGHLQTCRLLLSYGCDPNIISLQ[illegible] [illegible]EAA[illegible]QLLEAAFA 502

Qy        328 -DTE  -ICT --  --K[illegible]E-- [illegible]IKA[illegible]AKV[illegible]AVCKF[illegible] 372
Db        503 GDVE[illegible]VKKLCTVQSVNCRDIEGR[illegible] [illegible]AFDR[illegible]VP 562

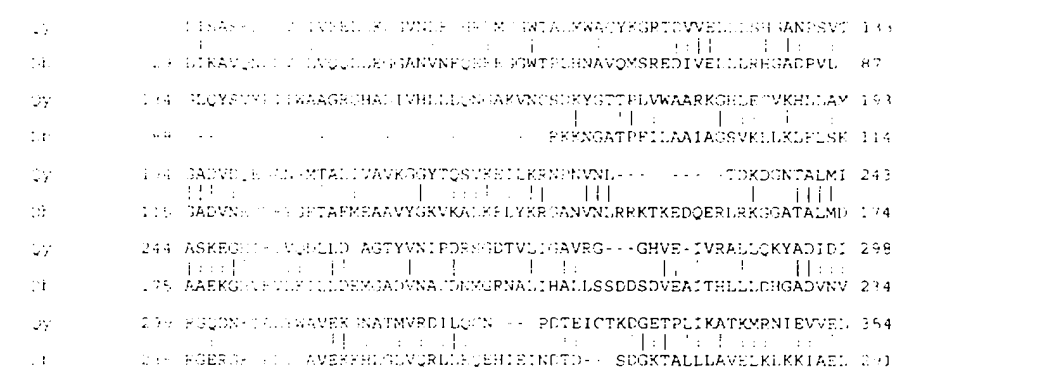
Qy        373 LHIAIRGHSRKLAELLLRN 391
Db        563 LHNA[illegible]SYGHYEVAELLVKH 581


RESULT 12
RN[illegible]A_HUMAN
ID   RN5A_HUMAN     STANDARD;      PRT;   [illegible]
AC   Q05823;
DT   01-NOV-1995 (Rel. 32, Created)
DT   [illegible]-FEB-1996 (Rel. 33, Last sequence [illegible])
DT   15-SEP-2003 (Rel. 42, Last annotation [illegible])
DE   2-5A-dependent ribonuclease (EC 3.[illegible]) [illegible] (2-5A-dependent [illegible])
DE   (Ribonuclease L) (RNase L) (Ribo[illegible]) [illegible]
GN   RNASEL OR RNS4
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. [illegible] AND MUTAGENESIS OF LYS-240 AND LYS-274.
RC   TISSUE=Kidney;
RX   MEDLINE=93201538; PubMed=7680958;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAase: a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:753-765(19[illegible]).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20515260; PubMed=11063255;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes: mediators
RT   of interferon action.";
RL   Mamm. Genome 11:989-992(2000).
RN   [3]
RP   CHARACTERIZATION [illegible] RNASEL ACTIVITY.
RX   MEDLINE=94245737; PubMed=7514601;
RA   Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
RA   Silverman R.H.;
RT   "Intrinsic molecular activities of the interferon-induced 2-5A-
RT   dependent RNase.";
RL   J. Biol. Chem. 2[illegible]:14153-14158(1994).
RN   [4]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11585831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in interferon alpha treated H9 cells.";
RL   J. Biol. Chem. 2[illegible]:48473-48482(2001).
RN   [5]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13354-13354(2002).
RN   [6]
RP   REVIEW.
RX   MEDLINE=[illegible]; PubMed=[illegible]56285;
RA   [illegible], W[illegible] F.A., Youle R.J.;
RT   "The 2-5A system in viral infection and apoptosis.";
RL   [illegible]
RN   [illegible]
RP   MUTAGENESIS OF [illegible].
RX   MEDLINE=99[illegible]; PubMed=[illegible]62963;
RA   Dong B., Silverman R.H.;
RT   "Alternative function of a protein kinase homology domain in 2',
RT   5'-oligoadenylate dependent RNase L.";
RL   Nucleic Acids Res. 27:439-445(1999).
RN   [8]
RP   MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND
RP   HIS-672.
RX   MEDLINE=21230542; PubMed=11333017;
RA   Dong B., Niwa M., Walter P., Silverman R.H.;
RT   "Basis for regulated RNA cleavage by functional analysis of RNase L
RT   and Ire1p.";
RL   RNA 7:361-373(2001).
RN   [9]
RP   VARIANTS SER-59; PHE-406; GLN-462 AND GLU-541.
RX   MEDLINE=21947464; PubMed=11941539;
RA   Rokman A., Ikonen T., Seppaelae E.H., Nupponen N., Autio V.,
RA   Mononen N., Bailey-Wilson J., Trent J., Carpten J., Matikainen M.P.,
RA   Koivisto P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
RT   "Germline alterations of the RNASEL gene, a candidate HPC1 gene at
RT   1q25, in patients and families with prostate cancer.";
RL   Am. J. Hum. Genet. [illegible]:1299-1304(2002).
RN   [10]
RP   VARIAN[illegible] SER-[illegible] [illegible] 462 AND GLU-541.
```

RX MEDLINE=21670932; PubMed=11179194.
RA Carpten J., Nupponen N., Isaacs S., Sood R., Robbins C., Xu J.,
RA Faruque M., Moses T., Ewing C., Gillanders E., Hu P., Panabieres K.,
RA Makalowska I., Baffoe-Bonnie A., Isaacs W., Smith J., Kainu T.,
RA Wiley K., Bronstein M., Ethier S., Montie J., Patterson B.,
RA Hostetter G., Matikainen M., Sood R., Kallioniemi O.-P., Trent J.,
RA Xiang Y., Wahl P., De Vere White R., Piantadosi S., Pretlow T.,
RA Burk R., Meyers D., Gronberg H., Walsh P., Kallioniemi A.,
RA Bailey-Wilson J., Walsh E., Isaacs W., Brown E.,
RT "Germline mutations in the ribonuclease L gene in families showing
RT linkage with HPC1.",
RL Nat. Genet. 27:181-184 (2001).
RN [11]
RP CHARACTERIZATION OF VARIANTS ASN-62; ASP-77 AND LEU-541.
RX MEDLINE=22045600; PubMed=12167674;
RA Casey G., Neville P.J., Plummer S.J., Xiang Y., Krumroy L.M.,
RA Klein E.A., Catalona W.J., Nupponen N., Carpten J.D., Trent J.M.,
RA Silverman R.H., White J.S.,
RT "RNASEL Arg462Gln variant is implicated in up to 13% of prostate
RT cancer cases.",
RL Nat. Genet. 32:581-583 (2002).
CC -!- FUNCTION: Endoribonuclease, mediates the interferon action, which
CC play a role in mediating resistance to virus infection and
CC apoptosis. Might play a role in the regulation of mRNA
CC turnover.
CC -!- CATALYTIC ACTIVITY: Cleaves single-stranded RNA, with preference for
CC UU and CA sequences, to give 5'-hydroxyl and 3'-phosphate termini
CC between 4 and 12 nucleotides in size.
CC -!- COFACTOR: Optimal RNA cleavage requires the presence of
CC either manganese or magnesium ions.
CC -!- ENZYME REGULATION: Activity is inhibited by 2'-5' oligoadenylates
CC linked cross-oligomerizes to form 2-5A, which is a potent
CC activator. Forms inhibited by ATP, GTP, inhibitors.
CC -!- SUBUNIT: Monomer (inactive) or homodimer (active).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (Probable).
CC -!- TISSUE SPECIFICITY: Highly expressed in placenta and kidney and
CC by prostate, testis, ovarian, small intestine, colon and peripheral
CC blood leukocytes.
CC -!- INDUCTION: By interferon.
CC -!- DOMAIN: The kinase domain at the C-terminus is inactive.
CC -!- DOMAIN: The nine ankyrin repeats are required to bind 2-5A with a
CC constitutive low 2-5A binding activity.
CC -!- DISEASE: Defects in RNASEL may be involved in susceptibility to
CC hereditary prostate cancer (HPC) [MIM:176807].
CC -!- SIMILARITY: Contains 2 ANK repeats.
CC ---- ----- -----
CC This SWISS-PROT entry is copyrighted. It is produced through a
CC collaboration between the Swiss Institute of Bioinformatics and
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for
CC commercial entities requires a commercial licence in some countries.
CC For information about licensing, see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch
CC ---- ----- -----
DR EMBL; L10381; AAA18072.1;
DR HSSP; P80144; 2MYC.
DR Genew; HGNC:10050; RNASEL.
DR MIM; 180435;
DR MIM; 601831;
DR GO; GO:0004521; Endoribonuclease activity; TAS.
DR GO; GO:0003723; RNA binding; IEA.
DR GO; GO:0006468; Protein amino acid phosphorylation; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000567; PUG.
DR Pfam; PF00023; ank; 9.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 8.
DR SMART; SM00056; PUG; 1.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc finger; Repeat;
KW ANK repeat; Polymorphism.
FT REPEAT 24 53 ANK 1.
FT REPEAT 58 87 ANK 2.
FT REPEAT 91 120 ANK 3.
FT REPEAT 124 153 ANK 4.
FT REPEAT 167 197 ANK 5.
FT REPEAT 201 234 ANK 6.
FT REPEAT 238 268 ANK 7.
FT REPEAT 272 301 ANK 8.
FT REPEAT 306 339 ANK 9.
FT DOMAIN 367 680 PROTEIN KINASE LIKE.
FT DOMAIN 58 641 RIBONUCLEASE.
FT DOMAIN 223 342 2-5A BINDING (P-LOOP) 1.
FT DOMAIN 254 276 2-5A BINDING (P-LOOP) 2.
FT ZN_FING 395 444 C6 TYPE (ATYPICAL).
FT VARIANT 59 59 G -> S.
FT /FTId=VAR_013509.
FT VARIANT 401 406 S -> P.
FT /FTId=VAR_013510.
FT VARIANT 462 462 R -> Q (risk factor for prostate cancer;
FT reduced enzymatic activity;
FT dbSNP:486907).
FT /FTId=VAR_012056.
FT VARIANT 541 541 S -> B (no change in enzymatic activity;
FT dbSNP:627918).
FT /FTId=VAR_012057.
FT MUTAGEN 24 240 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT ALMOST COMPLETE LOSS OF 2-5A BINDING
FT ACTIVITY WHEN ASSOCIATED WITH N-274.
FT MUTAGEN 274 274 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT ALMOST COMPLETE LOSS OF 2-5A BINDING
FT ACTIVITY WHEN ASSOCIATED WITH N 240.
FT MUTAGEN 392 392 K->R: COMPLETE LOSS OF ENZYMATIC ACTIVITY
FT AND ENZYME DIMERIZATION. NO CHANGE IN
FT BINDING TO 2-5A AND RNA.
FT MUTAGEN 561 561 H->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT MUTAGEN 564 564 R->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT MUTAGEN 632 632 W->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT MUTAGEN 661 661 D->A: COMPLETE LOSS OF ENZYMATIC
FT ACTIVITY.
FT MUTAGEN 667 667 R->A: COMPLETE LOSS OF ENZYMATIC
FT ACTIVITY. NO CHANGE IN 2-5A BINDING AND
FT ENZYME DIMERIZATION.
FT MUTAGEN 672 672 E->A: COMPLETE LOSS OF ENZYMATIC
FT ACTIVITY. NO CHANGE IN 2-5A BINDING



```
Db        292 LCKRGASTDC-- -GDLVV TAPRNYDRSGLVKYLLSHGAKPDPHPPAELWKPQSSHWGAA 346

Qy        333  KDGRELDYRPNKAGETPYNIDVSRLPSITLQIPGARHLSPTETDGDMLGY 441
               ||    :  |   :  ||          :  |          |:
Db        347 LKDLHPIYRP-MIGLKPPILEKYE          - IAPIGPGIYLGP 383


RESULT 13
ANR6_HUMAN
ID   ANR6_HUMAN      STANDARD;      PRT;   [illegible] AA.
AC   Q9Y2G4; Q9NU24; Q9UFQ9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat domain protein 6.
GN   ANKRD6 OR KIAA0957
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RX   MEDLINE=99244064; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [2]
RP   SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).
RC   TISSUE=Testis;
RA   Poustka A., Klein M., Mewes H. W.;
RL   Submitted (SEP 1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 140-692 FROM N.A. (ISOFORM 2).
RA   Tracey A.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         Isoid=Q9Y2G4-1; Sequence=Displayed;
CC       Name=2;
CC         Isoid=Q9Y2G4-2; Sequence=VSP_000270;
CC         Note=No experimental confirmation available.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- [illegible]
CC   -!- [illegible]
CC   -!- This SWISS-PROT entry is [illegible] between the Swiss Institute
CC   between the Swiss Institute [illegible] the European Bioinformatics
CC   the European Bioinformatics [illegible] use by non-profit institutes
CC   use by non-profit institutes [illegible] modified and this statement
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -!- [illegible]
DR   EMBL; AP023174; BAA76801.1;
DR   EMBL; AL117604; CAB55968.1; -.
DR   EMBL; AL096674; CAB86658.1; -.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:17280; ANKRD6.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 7.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 8.
DR   PROSITE; PS50297; ANK_REP_REGION.
KW   Repeat; ANK repeat; Alternative [illegible]
FT   REPEAT       9     38       ANK 1.
FT   REPEAT      41     70       ANK 2.
FT   REPEAT      74    103       ANK 3.
FT   REPEAT     107    136       ANK 4.
FT   REPEAT     140    169       ANK 5.
FT   REPEAT     173    202       ANK 6.
FT   REPEAT     206    235       ANK 7.
FT   REPEAT     239    268       ANK 8.
FT   VARSPLIC   264    264       Q -> QVLRPSRGRSLRKKRERLKEERRAQSVPRDEVAQGK
FT                               (in isoform 2).
FT                               /FTId=VSP_000270.
FT   CONFLICT    87     87       N -> I (IN REF. 2).
SQ   SEQUENCE   692 AA;  75720 MW;  D3211B487A8C839F CRC64;

  Query Match             4.0%;  Score 65;  DB 1;  Length 692;
  Best Local Similarity   23.4%;  Pred. No. 1.1e-10;
  Matches  117;  Conservative   67;  Mismatches  174;  Indels   40;  Gaps   10;

Qy        128 LMWACYP-PTGVELLLSHGANPSVTNHLQYSVYPLIWAAGRGHADIVHLLLQNGAKVNCS 187
              :  |  |:    |:  ||  |  ::  |:  |:  |:|  :|  :|  |  :
Db         14 LLVAASKQGTINVVQQLINKGARVAVT -KHGRTPLHLAANKGHLPVVQILLKAGCDLDVQ 72

Qy        188 DKYGTTPLVWAARKGHLECVKHLLAMGADVDQEGANSMTALIVAVKGGYTQSVKEILKRN 247
              |   : :  .  : |  :  |:  .  :|:  :  :  !  ::!'  :  :|
Db         73 DDGDQTALHWATVVGNTDIIAALIHHGCALDRQDKDGNTALHEASWHGFSQSAKLLVKAG 132

Qy        248 PNVNLTDYKGNTALMIASKEGRTEIVQQLLDAGTYVNIPDRSGDTVLIGAVRGGHVEIVP 307
                 ||   :|  :  |:  :  |   |:  :: :  :|:  | |  |  |:  |:
Db        133 ANVLARNKAGNTALHLACQNSHSQSTEVLLLAGSHADLKNNAGDTCLHVAARYNHLSIIR 191

Qy        308 ALLQKYASITPDRGQDNKTALYWAVEKNATMVRDILQCNPDTEICTKDGETPLIKATKMR 367
              ||  :  :  |   |||.      :  :  :  :    :  |      :|||  '
Db        192 LLLTAPNVHLKNQAGDTALHVAAALNRKKVAKILLEAGADTTIVNNAGQTPLETARYHN 251

Qy        368 NIEVVELLIDRGAKVSAVDKKGDTPLHIAIRGRSEKLAELLLRNPK - --DGRLLYRP 424
              |  |     |  :  :  |  |  :  :  :  |  |  :::      |  |
Db        252 NPEVA [illegible] SAGDTPSSEQAVARKEEAREEPDSASPEPRAKCDRRRKSRF 308

Qy        425 NKAG-   ETPYNIDPSHQKSILTQIPGARHLSPTETD----------GDMLGYDLYSSA 473
                  :  ||':  :  |  :  :  :  :  :  :  :  |  |  :
Db        309 KVSAPSIPIPPAPQQQPGHQRNL----- -HAHNHPKKRNRHRCSSPPPPHEPRAYQLYT-- 360

Qy        474 LADILSEFINQPPICVGLYAQWDSGKSFLLKKLEDEMK 511
              |       ||  .       |     |  ||::::
Db        361 LYRGKDGRVMQAPIN-  --  GCRCEPLINKLENQLE 391


RESULT 14
ANR3_HUMAN
ID   ANR3_HUMAN      STANDARD;      PRT;   832 AA.
AC   P57078; Q96KH0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat
DE   domain-containing protein 3) (IBP-delta interacting protein kinase).
GN   ANKRD3 OR DIK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal kidney, and Fetal lung;
RA   Shimizu N., Kubota T., Shibuya K.;
RL   Submitted (AUG 2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289744; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schon O., Desario A., Reichelt J., Kauer G., Blocker H.,
```

```
RA   Ramsey C., Beck A., Klages S., Hennig S., Bressel M., Tajbakhsh S.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nicholson D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome [illegible]";
RL   Nature 425:311-319 (2003).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; [illegible]
CC       Name=1;
CC         IsoId=[illegible]; [illegible]
CC       Name=2;
CC         IsoId=[illegible]; [illegible]
CC   -!- SIMILARITY: BELONGS TO THE [illegible] FAMILY. [illegible]
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement. See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch.
CC   -----------------------------------------------------------------------
DR   EMBL; AB047783; BAB56146.1; -.
DR   EMBL; AB001743; BAA95526.1; -.
DR   HSSP; P25963; 1IKN.
DR   Genew, HGNC:448; ANKRD4.
DR   MIM; 604706; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001245; Ser_thr_pk_1.
DR   InterPro; IPR002290; Ser_thr_pk_2.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00023; ank; 10.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
KW   Transferase; Serine/threonine-protein [illegible]; ATP-binding; Repeat;
KW   ANK repeat; Alternative splicing.
FT   DOMAIN       22    286       PROTEIN KINASE.
FT   REPEAT      485    514       ANK 1.
FT   REPEAT      518    547       ANK 2.
FT   REPEAT      551    580       ANK 3.
FT   REPEAT      584    613       ANK 4.
FT   REPEAT      617    646       ANK 5.
FT   REPEAT      650    679       ANK 6.
FT   REPEAT      683    712       ANK 7.
FT   REPEAT      716    745       ANK 8.
FT   REPEAT      749    780       ANK 9.
FT   REPEAT      783    811       ANK 10.
FT   NP_BIND      14     36       ATP (POTENTIAL).
FT   BINDING      51     51       ATP (BY SIMILARITY).
FT   ACT_SITE    143    143       BY SIMILARITY.
FT   VARSPLIC    309    325       Missing (in isoform 2).
FT                                /FTId=VSP_001743.
FT   CONFLICT    714    714       M -> T (IN REF. 1).
SQ   SEQUENCE   811 AA;  91910 MW;  5[illegible] CRC32;

  Query Match              [illegible]%;  Score [illegible];  [illegible];
  Best Local Similarity    27.9%;  Pred. No. [illegible];
  Matches   99;  Conservative   94;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          35 NEDQQTPLMIAAEQGNLEIVKEL[illegible]  [illegible]  94

Db         [illegible] NRRGSTPLHMAVERRVEVVELL[illegible]  [illegible]

Qy          95 VNLEHRDMG[illegible]  [illegible]  154

Db         576 ASVNEVDPERTPMHVACQHGQENIVRILLRRGVDVSLQG-KDAWLPLHYAAWQGHLPIV 634

Qy         155 HLLL-QNSAKVNCSEKYGTTPLAWAARFGHLECVKHLLAMGADVDQEGANSMTALIVAVK 213

Db         635 KLLAKQTNGVNAQTLDGKTPLHLAAQFGHYRVARILIDLGSDVNVCSLLAQTPLHVAAE 694

Qy         214 DGYTQNVPEIIKPRNPNVNLTDKEGNTALMIASKEGHTRIVQDLLDAGTYVNIPDRSGDTV 273

Db         695 DGRTSIA[illegible]HRGASKEAMDSDGYTALHLAA  -  [illegible]  726

Qy         274 LICAVRTGRVEIVRALIQKFYADIDIRNQDNKTALYWAVEKGNATMVRDILQCNPDTRICT 333

Db         727 --  RDHRATPVKLLVEEKADVLAKDPLNQTALHLAAAHGHSEVVEELV[illegible]  771

Qy         334 KMGETPLHKATKNRNIEVVELLLKFGAKVSAVDKKGDTPLHIAIRGRSPKLARLLLRN 391

Db         772 --  --         GSATVIDLP      --  DEQGLSALHLAAQGRHAQTVETLLSH 805


RESULT 15
RN5A_MOUSE
ID   RN5A_MOUSE     STANDARD;      PRT;   735 AA.
AC   Q05921; Q9ERU7;
DT   01-NOV-1995 (Rel. 32, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   2',5'-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN   RNASEL OR RNS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3H/An; TISSUE=Adipose tissue;
RX   MEDLINE=95051629; PubMed=7961756;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and mapping of the human and mouse RNase L genes: mediators
RT   of interferon [illegible].";
RL   Mamm. Genome 6:[illegible] (1995).
RN   [2]
RP   SEQUENCE OF 1-[illegible] FROM N.A.
RX   MEDLINE=92205179; PubMed=1680959;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAase, a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:[illegible] (1993).
RN   [3]
RP   FUNCTION, AND [illegible] (RNASE L ACTIVITY).
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNA stability in interferon alpha-treated H9 cells.";
RL   J. Biol. Chem. [illegible]:[illegible] (2001).
RN   [4]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. [illegible]:[illegible] (2002).
CC   -!- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of mRNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 3' of UpNp dimers, with preference for
CC       UU and UA sequences, to sets of discrete products ranging from
CC       between 4 and 12 nucleotides in length.
CC   -!- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC       either manganese or magnesium and ATP.
CC   -!- ENZYME REGULATION: After binding to 2-5A (5' phosphorylated 2',5'
CC       linked oligoadenylates) the homodimerization and subsequent
CC       activation [illegible] inhibited by RNase L inhibitor.
```

CC -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC -!- TISSUE SPECIFICITY: Expressed in spleen, thymus, lung, testis,
CC kidney, liver and heart.
CC ! INDUCTION: By interferons.
CC -!- DOMAIN: A single active ribonuclease domain in a dimer is
CC sufficient for ribonuclease activity.
CC ! DOMAIN: The kinase domain allows the autophosphorylation.
CC -!- DOMAIN: The nine ankyrin repeats all, called 'A' here,
CC constitute the 2-5A binding domain.
CC -!- SIMILARITY: Contains 9 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF281045; AAG33708.1; -.
DR EMBL; L13382; AAA36117.1; -.
DR PIR; B45771; B45771.
DR HSSP; P42773; 1IH5.
DR MGD; MGI:1098272; Rnasel.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000567; PUG.
DR Pfam; PF00023; ank; 8.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01415; ANKYRIN.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 8.
DR SMART; SM00580; PUG; 1.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hydrolase; Nuclease; RNA-binding; Kinase; Zinc-finger;
KW Repeat; ANK repeat.
FT REPEAT 24 53 ANK 1
FT REPEAT 58 87 ANK 2
FT REPEAT 91 120 ANK 3
FT REPEAT 124 153 ANK 4
FT REPEAT 167 197 ANK 5
FT REPEAT 201 234 ANK 6
FT REPEAT 238 268 ANK 7
FT REPEAT 272 301 ANK 8
FT REPEAT 303 [illegible] ANK 9
FT DOMAIN [illegible] PK_TYR_KIN
FT DOMAIN [illegible] PUG
FT DOMAIN [illegible] [illegible]
FT DOMAIN [illegible] [illegible]
FT ZN_FING 401 436 [illegible]
SQ SEQUENCE 735 AA; 83174 MW; [illegible]

Query Match 3.9%; Score 46%; [illegible]
Best Local Similarity 32.5%; Pred. No. [illegible]
Matches 111; Conservative 76; Mismatches 112; Indels 44; Gaps 11;

Qy 74 TALISASKEGHVHIVEELLKAGVN [illegible] 131
Db 27 SSLIKAVQKGDVVRVQQLLEKGAD [illegible] 85

Qy 132 VTGLQYSVYPIIWAAAGRGHADIVR [illegible] 191
Db 86 -- -- - --- --- -- [illegible] 112

Qy 192 AMGADVDQEGANSMTALIVAVKG [illegible] 241
Db 113 SCGADVNECDENGFTAFMEAAER [illegible] 172

Qy 242 MIASKEGHTEIVQLLL PASTYVS [illegible] 296

Db 173 MSAAEKGHLEVLRILLNDMKAEVDARDNMGRNALIRTLLNWDCENVEEITSILIQHGADV 232

Qy 297 DIRGQDNKTALYWAVEKGNATMVRDIL--CNPDTEICTKDGETPLIKATKMRNIEVVE 353
Db 233 NVRGERGKTPLIAAVERKETGLVQMLLSREGININDAR--DNEGKTALLIAVDKQLKEIVQ 290

Qy 354 LLLDKGAKVSAVDKKGQTPLHIAIRGRSRKLAELLL---RNP 392
Db 291 LLLEKGA DKQDDLVWIARRNHDYHLVKLLLPYVANP 326

Search completed: November 21, 2003, 12:49:27
Job time : 32.5 secs